1981

Optimality conditions and dual formulations for programming problems over cone domains

Bradley Owen Skarpness
Iowa State University

8128861

SKARPNESS, BRADLEY OWEN

OPTIMALITY CONDITIONS AND DUAL FORMULATIONS FOR PROGRAMMING PROBLEMS OVER CONE DOMAINS

Iowa State University

PH.D. 1981

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

Optimality conditions and dual formulations for

programming problems over cone domains

by

Bradley Owen Skarpness

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Co-Majors: Statistics
Industrial Engineering

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1981

TABLE OF CONTENTS

1. INTRODUCTION

Consider the general mathematical programming problem:

$$\text{minimize (maximize)} \quad f(x)$$
$$\text{subject to} \quad g_i(x) \leq b_i \; ; \; i = 1, 2, \ldots, m ,$$

where $f(x)$ and $\{g_i(x)\}$ are numerical valued functions of x and the b_i's are known constants. Special cases of this problem are linear programming problems in which $f(x)$ and $\{g_i(x)\}$ are linear functions of x, as well as convex programming problems in which $f(x)$ and $\{g_i(x)\}$ are convex (concave) functions of x. This second type of problem is a special case of a more general class of programming problems known as nonlinear programming problems in which $f(x)$ and/or $\{g_i(x)\}$ are nonlinear functions of x.

The area of linear programming has been extensively researched since George Dantzig's development in 1947 of the Simplex algorithm to solve linear programming problems.

Kuhn and Tucker (1951) addressed themselves to nonlinear programming problems. They considered the special case in which $f(x)$ and $\{g_i(x)\}$ were required to be convex (concave) differentiable functions for nonnegative x. They also introduced a set of conditions which are known as the Kuhn-Tucker conditions and showed that solving the minimization (maximization) programming problem was equivalent to finding a saddle value solution for a certain Lagrangian function which was in turn equivalent to satisfying the Kuhn-Tucker conditions.

Their work has led to the area of duality through the saddle value problem and the Kuhn-Tucker conditions. That is, associated with a nonlinear programming problem is another problem which, when solved, yields the optimal solution to the original problem, and conversely.

Since Kuhn and Tucker's initial investigation, the area of nonlinear programming has expanded in several directions, including many types of programming models.

It has been shown, Mangasarian (1969), that in the case of convex programming not all the properties of convex and concave functions are needed to prove, say, the equivalence between a minimization programming problem and the Kuhn-Tucker conditions. Thus, some of the results of Kuhn and Tucker have been extended to a larger class of functions known as quasiconvex (quasiconcave) or pseudoconvex (pseudoconcave) functions.

Quasiconvex functions were first mentioned by Nikaidô (1954), and later, Tuy (1964) was the first to introduce functions which are pseudoconvex. In Chapter 2 we define this larger class of convex (concave) numerical functions known as quasiconvex (quasiconcave) and pseudoconvex (pseudoconcave) functions. Other definitions of similar functions are defined and examples as well as counter-examples are given to illustrate the similarities and differences between these types of functions.

Using these generalized functions, we conclude Chapter 2 with a result which extends Bhatt and Misra (1975) sufficient conditions for optimality of a Fritz John stationary point problem.

That is, find $\bar{x} \varepsilon P \subset E^n$, $\bar{r}_o \varepsilon E^1$, $\bar{r} \varepsilon E^m$, if they exists, such that

$$\bar{r}_o \nabla f(\bar{x}) + \bar{r}'\nabla(g(\bar{x})-b) = 0$$

$$g(\bar{x}) - b \leq 0$$

$$\bar{r}'(g(\bar{x})-b) = 0$$

$$(\bar{r}_o,\bar{r}) \geq 0 .$$

In Chapter 3, we consider a modified Kuhn-Tucker stationary point problem:

find an $\bar{x} \varepsilon P \subset E^n$ and $\bar{u} \varepsilon -C^* \subset E^m$ such that,

$$\nabla' f(\bar{x}) + \bar{u}'\nabla(g(\bar{x})-b) = 0$$

$$\bar{u}'(g(\bar{x})-b) = 0$$

$$g(\bar{x}) - b \varepsilon C$$

where C is an arbitrary cone in E^m . We establish necessary and sufficient conditions between this problem and a certain class of nonlinear programming problems where the constraints are in arbitrary cone domains.

These results are applied to two problems. The first is a modified Farkas Lemma over degenerate and nondegenerate cone domains which was established using only a "partial" linear duality theorem. The second, a quadratic programming problem over cone domains in which strong duality results are established between the original problem and its dual.

The necessary and sufficient optimality conditions presented in Chapter 3 are developed for problems whose constraints are linear functions over arbitrary cone domains. This structure allows us to consider problems with nonlinear constraints over equality and inequality restrictions which are considered in the usual nonlinear programming setting. Moreover, this structure allows us to establish necessary and sufficient conditions even though the problem is basically nonlinear in structure. These conditions are established by imposing a certain rank condition rather than the usual constraint qualification or interior conditions, see Mangasarian (1969).

These optimality conditions are subsequently used in Chapter 4 to generate more than one dual problem. These problems are developed for linear fractional problems over cone domains; i.e. problems where the objective function is composed of the quotient of two linear functions. These dual problems differ in structure from the classical formulation presented by Charnes and Cooper (1962). A diagram is given showing the relationship between the results which are established.

Alders (1976) established necessary and sufficient conditions for certain types of nonlinear programming problems with nonlinear constraints over arbitrary cone domains. The avenue explored in this thesis considers nonlinear programming problems with linear constraints over cone domains; this structure allows us to develop dual problems with degenerate as well as nondegenerate cone domains in the classical spirit by appealing to a certain rank condition imposed on the constraints.

2. GENERALIZED CONVEX FUNCTIONS

2.1. Introduction

Optimality condition and dual formulations of programming problems, to a great extent, rely upon the class of functions involved. In subsequent sections we define various types of functions, give some of their salient properties, and consider relationships between them. Our presentation follows that of Mangasarian (1969).

These properties and relationships are used to establish in the presence of equality-inequality constraints, a sufficient optimality criteria of the Fritz John type for certain nonlinear programming problems.

2.2. Definitions and Properties of Generalized Convex Functions

This section presents the definitions of quasiconvex (quasiconcave), strictly-quasiconvex (strictly-quasiconcave), pseudoconvex (pseudoconcave), convex (concave), and strictly convex (strictly concave) function. These functions are initially defined for a point $\bar{x} \in P \subset E^n$, where E^n is the Euclidean space of dimension n. If the definition holds for each point in P, then we say the function is quasiconvex, etc., on P. The set $P \subset E^n$ is the set on which the functions are defined. If $\theta(x)$ is a function defined on some open set $P \subset E^n$; i.e. $\theta: E^n \rightarrow E^1$, we will denote $\nabla\theta(\bar{x})$ as the n-dimensional gradient vector of θ at $\bar{x}$, that is $\nabla\theta(\bar{x}) = (\theta_1, \theta_2, \ldots, \theta_n)'$ where θ_i is the partial derivative of θ with respect to x_i, evaluated at $\bar{x} = (\bar{x}_1, \bar{x}_2, \ldots, \bar{x}_n)'$. Furthermore, if $g(x)$ is an m-dimensional vector function defined on an open set $P \subset E^n$; i.e. $g: E^n \to E^m$, then $\nabla g(\bar{x})$ will denote the $m \times n$ Jacobian matrix of first order partial derivatives, that is $\partial g_i(\bar{x})/\partial x_j$ $(i = 1,2, \ldots, m;\ j = 1,2, \ldots, n)$ evaluated at $\bar{x}$. We shall denote transpose with a prime, that is, if $\bar{r}$ is a column vector in E^m then $\bar{r}'$ will be its corresponding row vector.

<u>Definition 2.2.1.</u> Let θ be defined on an open set $P \subset E^n$, and let $\bar{x} \in P$. θ is said to be differentiable at $\bar{x}$ if for all $x \in E^n$ such that $\bar{x} + x \in P$ we have

$$\theta(\bar{x}+x) = \theta(\bar{x}) + t(\bar{x})x + \alpha(\bar{x},x)\ ||x||$$

where $t(\bar{x})$ is an n-dimensional bounded vector, and α is a numerical function of x such that $\lim_{x \to 0} \alpha(\bar{x},x) = 0$.

Let θ be defined on an open set $P \subset E^n$, and let $\bar{x} \in P$, it can be shown, using Definition 2.2.1, that:

(i) If θ is differentiable at $\bar{x}$, then θ is continuous at $\bar{x}$, and $\nabla\theta(\bar{x})$ exists (but not conversely), and

$$\theta(\bar{x}+x) = \theta(\bar{x}) + \nabla'\theta(\bar{x})x + \alpha(\bar{x},x)\ ||x||$$

and $\lim_{x \to 0} \alpha(\bar{x},x) = 0$ for $\bar{x} + x \in P$.

(ii) If $\nabla\theta(\bar{x})$ exists and $\nabla\theta$ is continuous at $\bar{x}$, then θ is differentiable at $\bar{x}$.

<u>Definition 2.2.2.</u> A numerical function θ defined on $P \subset E^n$ is said to be quasiconvex (QCX) at $\bar{x} \in P$ if for each $x \in P$ such that

$$\theta(x) \leqq \theta(\bar{x}) \quad \text{implies} \quad \theta[(1-\lambda)\bar{x} + \lambda x] \leqq \theta(\bar{x})$$

where $0 \leqq \lambda \leqq 1$ and $(1-\lambda)\bar{x} + \lambda x \in P$.

Definition 2.2.3. A numerical function θ defined on $P \subset E^n$ is said to be quasiconcave (QCA) at $\bar{x} \in P$ if for each $x \in P$ such that

$$\theta(x) \geqq \theta(\bar{x}) \quad \text{implies} \quad \theta[(1-\lambda)\bar{x} + \lambda x] \geqq \theta(\bar{x})$$

where $0 \leqq \lambda \leqq 1$ and $(1-\lambda)\bar{x} + \lambda x \in P$.

We should note here that if θ is quasiconvex at $\bar{x} \in P$ if and only if $-\theta$ is quasiconcave at $\bar{x} \in P$. Results obtained for quasiconvex functions can be changed into results for quasiconcave functions by appropriate multiplication by -1, and vice versa.

Theorem 2.2.1. Let θ be a numerical function defined on a convex set $P \subset E^n$ and let

$$\Lambda_\alpha = \{x : x \in P,\ \theta(x) \leqq \alpha\}$$

$$(\Omega_\alpha = \{x : x \in P,\ \theta(x) \geqq \alpha\}) ,$$

then θ is quasiconvex (quasiconcave) on P if an only if $\Lambda_\alpha(\Omega_\alpha)$ is convex for each $\alpha \in E^1$.

Proof: We prove only the quasiconvex case. ($\Rightarrow$) Let θ be QCX on P, $\alpha \in E^1$, and $x^1, x^2 \in \Lambda_\alpha$. Let $\theta(x^2) \leqq \theta(x^1)$ then since $x^1, x^2 \in \Lambda_\alpha$ we have $\theta(x^2) \leqq \theta(x^1) \leqq \alpha$. Since θ is QCX, and since P is convex, we have that for $0 \leqq \lambda \leqq 1$,

$$\theta[(1-\lambda)x^1 + \lambda x^2] \leqq \theta(x^1) \leqq \alpha .$$

Hence, $(1-\lambda)x^1 + \lambda x^2 \in \Lambda_\alpha$, and Λ_α is convex. ($\Leftarrow$) Let $x^1, x^2 \in P$, $\theta(x^2) \leqq \theta(x^1)$, and $0 \leqq \lambda \leqq 1$. If we let $\alpha = \theta(x^1)$, then since Λ_α is convex we have that

$$\theta[(1-\lambda)x^1 + \lambda x^2] \leqq \alpha = \theta(x^1)$$

and hence, θ is quasiconvex on P.

The next theorem gives a necessary and sufficient condition for a differentiable function θ to be QCX (QCA). Again, we omit the QCA proof, since it is similar to the QCX proof.

<u>Theorem 2.2.2.</u> Let P be an open set in E^n, and let θ be a numerical function defined on P. Then for $x^1, x^2 \in P$

$$\left.\begin{array}{c}\theta \text{ differentiable and} \\ \text{QCX at } x^1\end{array}\right\rangle \Rightarrow \left\langle \theta(x^2) \leqq \theta(x^1) \Rightarrow \nabla'\theta(x^1)(x^2-x^1) \leqq 0 \right\rangle$$

$$\theta \text{ QCX on } P \Leftarrow \left\langle \begin{array}{l} P \text{ convex, } \theta \text{ differentiable on } P \\ \left\langle \theta(x^2) \leqq \theta(x^1) \Rightarrow \nabla'\theta(x^1)(x^2-x^1) \leqq 0 \right\rangle \end{array}\right.$$

$$\left.\begin{array}{c}\theta \text{ differentiable and} \\ \text{QCA at } x^1\end{array}\right\rangle \Rightarrow \left\langle \theta(x^2) \geqq \theta(x^1) \Rightarrow \nabla'\theta(x^1)(x^2-x^1) \geqq 0 \right\rangle$$

$$\theta \text{ QCA on } P \Leftarrow \left\langle \begin{array}{l} P \text{ convex, } \theta \text{ differentiable on } P \\ \left\langle \theta(x^2) \geqq \theta(x^1) \Rightarrow \nabla'\theta(x^1)(x^2-x^1) \geqq 0 \right\rangle \end{array}\right.$$

<u>Proof:</u> ($\Rightarrow$) If $x^1 = x^2$, the proof is trivial. Assume $x^1 \neq x^2$. Since P is open there exists an open ball $B_\delta(x^1)$ around x^1 with radius $\delta > 0$ contained in P. For $0 < \tilde{u} < 1$ and $\tilde{u} < \dfrac{\delta}{||x^2-x^1||}$ we have $\tilde{x} = x^1 + \tilde{u}(x^2-x^1) = (1-\tilde{u})x^1 + \tilde{u}x^2 \in B_\delta(x^1)$. If

$\theta(x^2) \leqq \theta(x^1)$ then $\theta(\tilde{x}) \leqq \theta(x^1)$ since θ is QCX. Therefore,

$$\theta[(1-\lambda)x^1 + \lambda\tilde{x}] \leqq \theta(x^1) \tag{2.2.1}$$

because θ is QCX and B_δ is convex.

Now using the fact that if θ is differentiable at a point $\bar{x}$, we have $\theta(\bar{x}+x) = \theta(\bar{x}) + \nabla'\theta(\bar{x})x + \alpha(\bar{x},x)\ ||x||$. Letting $x^1 = \bar{x}$, $x = \lambda(\tilde{x}-x^1)$, and $\bar{x} + x = (1-\lambda)x^1 + \lambda\tilde{x}$, substituting into (2.2.1) we have

$$\lambda\nabla'\theta(x^1)(\tilde{x}-x^1) + \alpha[x^1,\lambda(\tilde{x}-x^1)]\ \lambda||\tilde{x}-x^1|| \leqq 0 \ .$$

Hence,

$$\nabla'\theta(x^1)(\tilde{x}-x^1) + \alpha[x^1,\lambda(\tilde{x}-x^1)]\ ||\tilde{x}-x^1|| \leqq 0, \ (0<\lambda<1)$$

$$\nabla'\theta(x^1)(\tilde{x}-x^1) \leqq 0 \quad (\text{by letting } \lambda \to 0)$$

$$\nabla'\theta(x^1)(x^2-x^1) \leqq 0 \ . \quad (\text{letting } \tilde{x} - x^1 = \tilde{u}(x^2-x^1) \text{ and } \tilde{u} > 0)$$

($\Leftarrow$) Let x^1, $x^2 \in P$, and $\theta(x^2) \leqq \theta(x^1)$, let $(x^1,x^2) = \{x : x = (1-\lambda)x^2 + \lambda x^1,\ 0 < \lambda < 1\}$ and let $\Omega = \{x : \theta(x^1) < \theta(x),\ x \in (x^1,x^2)\}$. Now if we can show $\Omega = \phi$ then it follows that θ is QCX. We assume that there is an $\bar{x} \in \Omega$ and show that a contradiction ensues. Since $\theta(x^2) \leqq \theta(x^1) < \theta(x)$, $x \in \Omega$, by hypothesis we have

$$\nabla'\theta(x)(x^1-x) \leqq 0 \ ,$$

and

$$\nabla'\theta(x)(x^2-x) \leqq 0 \ .$$

Since $x = (1-\lambda)x^1 + \lambda x^2$ then $x^2 - x = -\lambda(x^2-x^1)$ and $x^2 - x = (1-\lambda)(x^2-x^1)$ we have

$$\left.\begin{array}{c} -\lambda\nabla'\theta(x)(x^2-x^1) \leqq 0 \\ \text{and} \\ (1-\lambda)\nabla'\theta(x)(x^2-x^1) \leqq 0 \end{array}\right\rangle \text{implies } \nabla'\theta(x)(x^2-x^1) = 0$$

for $x \in \Omega$ and $0 < \lambda < 1$. Since $\theta(x^1) < \theta(\bar{x})$, and since θ is continuous on P (θ differentiable), the set Ω is open relative to (x^1,x^2), it contains $\bar{x}$, and there exists an $x^3 = (1-u)\bar{x} + ux^1$, $0 < u \leqq 1$, such that x^3 is a vector such that $\theta(x^3) = \theta(x^1)$. [The set Ω is open relative to (x^1,x^2) by the equivalent condition for continuous θ : that is the set $\{x : x \in P, \theta(x) > \alpha\}$ and $\{x : x \in P, \theta(x) < \alpha\}$ are open relative to P for each real α, let $P = (x^1,x^2)$, $\alpha = \theta(x^1)$].

By the Mean-Value Theorem [if θ is differentiable on an open convex set P, with x^1, $x^2 \in P$, then

$$\theta(x^2) - \theta(x^1) = \nabla'\theta(x^1 + \lambda(x^2-x^1))(x^2-x^1),\ 0 \leqq \lambda \leqq 1]$$

we have for some $\hat{x} \in \Omega$,

$$0 < \theta(\bar{x}) - \theta(x^1) = \theta(\bar{x}) - \theta(x^3) = \nabla'\theta(\hat{x})(\bar{x}-x^3)$$

$$= u\nabla'\theta(\hat{x})(\bar{x}-x^1)\ .$$

However, since $\bar{x} = (1-\bar{\lambda})x^1 + \bar{\lambda}x^2$, for some $\bar{\lambda} \in (0,1)$, then $0 < u\nabla'\theta(\hat{x})(\bar{x}-x^1) = u\bar{\lambda}\nabla'\theta(\hat{x})(x^2-x^1)$, for some $\bar{\lambda} > 0$, $u > 0$. Since $\hat{x} \in \Omega$, the last relation above contradicts the equality $0 = \nabla'\theta(\bar{x})(x^2-x^1)$ for all $\bar{x} \in \Omega$, which was established earlier. Hence, the result follows.

<u>Definition 2.2.4.</u> A numerical function θ defined on $P \subset E^n$ is said to be strictly-quasiconvex (SQCX) at $\bar{x} \in P$ if for each $x \in P$ such that

$$\theta(x) < \theta(\bar{x}) \quad \text{implies} \quad \theta[(1-\lambda)\bar{x} + \lambda x] < \theta(\bar{x})$$

where $0 < \lambda < 1$ and $(1-\lambda)\bar{x} + \lambda x \in P$.

<u>Definition 2.2.5.</u> A numerical function θ defined on $P \subset E^n$ is said to be strictly-quasiconcave (SQCA) at $\bar{x} \in P$ if for each $x \in P$ such that

$$\theta(x) > \theta(\bar{x}) \quad \text{implies} \quad \theta[(1-\lambda)\bar{x} + \lambda x] > \theta(\bar{x})$$

where $0 < \lambda < 1$ and $(1-\lambda)\bar{x} + \lambda x \in P$.

Obviously, θ is strictly-quasiconvex at $\bar{x} \in P$ if and only if $-\theta$ is strictly-quasiconcave at $\bar{x} \in P$.

In considering the relationship between functions which are SQCX and QCX we observe that a SQCX function need not be QCX. Consider the numerical function θ defined on E^1 as follows,

$$\theta(x) = \begin{cases} 1 \text{ for } x = 0 \\ 0 \text{ for } x \neq 0 \end{cases}$$

This function is SQCX at 0 but is not QCX at 0 . In particular taking $x^1 = -1$, $x^2 = 1$, $\lambda = 1/2$, we see that $\theta(x^2) = \theta(x^1)$, but $\theta[(1-\lambda)x^1 + \lambda x^2] > \theta(x^1)$.

We also observe that a QCX function on E^1 need not be SQCX on E^1 . Consider the numerical function θ defined on E^1 as follows,

$$\theta(x) = \begin{cases} x \text{ for } x \leqq 0 \\ 0 \text{ for } 0 < x < 1 \\ x-1 \text{ for } x \geqq 1 \end{cases}$$

This function is QCX on E^1 but not SQCX on E^1. For by taking $x^1 = 1/2$, $x^2 = -1/2$, $\lambda = 1/10$, then $\theta(x^2) < \theta(x^1)$, but $\theta[(1-\lambda)x^1 + \lambda x^2] = \theta(x^1)$, which contradicts Definition 2.2.4.

A SQCX (SQCA) function is essentially a restriction of QCX (QCA) function. We have shown that a SQCX function is not necessarily QCX, if however, we require θ to be lower (upper) semicontinuous, the above counter example will be eliminated and every SQCX (SQCA) function will also be QCX (QCA).

<u>Definition 2.2.6.</u> θ is lower semicontinuous at $\bar{x}$ if an only if

(i) given $\varepsilon > 0$, there exists $\delta > 0$, for all $x \in P$, such that, if $||x-\bar{x}|| < \delta$ then $-\varepsilon < \theta(x) - \theta(\bar{x}) < \varepsilon$ or equivalently

(ii) for all $x_m \to \bar{x}$, $\liminf\limits_{m\to\infty} \theta(x_m) \geqq \theta(\lim\limits_{m\to\infty} x_m) = \theta(\bar{x})$.

<u>Theorem 2.2.3.</u> Let θ be a lower (upper) semicontinuous numerical function defined on the convex set $P \subset E^n$. If θ is SQCX (SQCA) on P, then θ is QCX (QCA) on P.

<u>Proof:</u> We prove only the QCX case. Let θ be SQCX on P, with x^1 and $x^2 \in P$. By the definition of θ being SQCX we have if

$$\theta(x^2) < \theta(x^1) \text{ implies } \theta[(1-\lambda)x^1 + \lambda x^2] < \theta(x^1)$$

for $0 < \lambda < 1$.

If $\theta(x^2) < \theta(x^1)$ we are done. Assume $\theta(x^2) = \theta(x^1)$ and we shall show that there does not exists a $\hat{x} \varepsilon (x^1,x^2)$ such that $\theta(x^1) < \theta(\hat{x})$. [This states that $\theta(x^1) \geqq \theta(x^2)$ which implies θ is QCX.]

Assume there does exists $\hat{x} \varepsilon (x^1,x^2)$ such that $\theta(x^1) < \theta(\hat{x})$. Then

$$\hat{x} \varepsilon \Omega = \{x : \theta(x^1) < \theta(x),\ x \varepsilon (x^1,x^2)\} .$$

Since θ is lower semicontinuous on P, Ω is open relative to (x^1,x^2). Hence, there exists $\tilde{x} \varepsilon (x^1,x^2) \cap \Omega$. θ being SQCX and $\hat{x},\tilde{x} \varepsilon \Omega$ we have if

$$\theta(x^1) < \theta(\hat{x}) \quad \text{then} \quad \theta(\tilde{x}) < \theta(\hat{x}) \ [\tilde{x} \varepsilon (x^1,\hat{x})] , \tag{2.2.2}$$

$$\theta(x^2) < \theta(\tilde{x}) \quad \text{then} \quad \theta(\hat{x}) < \theta(\tilde{x}) \ [\hat{x} \varepsilon (\tilde{x},x^2)] \tag{2.2.3}$$

(2.2.2) and (2.2.3) yield a contradiction. Hence, no such $\hat{x}$ exists and θ is QCX on P.

An important property of SQCX (SQCA) functions is given by the next theorem.

Theorem 2.2.4. Let θ be a numerical function defined on the convex set $P \subset E^n$, and let $\bar{x} \varepsilon P$ be a local minimum (maximum). If θ is SQCX (SQCA) at $\bar{x}$, then $\theta(\bar{x})$ is a global minimum (maximum) of θ on P.

Proof: We give the proof for SQCX functions. Let $\bar{x}$ be a local minimum, then there exists $B_\delta(\bar{x})$ such that $x \varepsilon B_\delta(\bar{x}) \cap P$ implies

$\theta(\bar{x}) \leqq \theta(x)$. Assume there exists $\hat{x} \varepsilon P$, $\hat{x} \notin B_\delta(\bar{x})$ such that $\theta(\hat{x}) < \theta(\bar{x})$. Since θ is SQCX at $\bar{x}$ and P convex, we have

$$\theta[(1-\lambda)\bar{x} + \lambda\hat{x}] < \theta(\bar{x}) \tag{2.2.4}$$

for any $\lambda \varepsilon (0,1)$. But for $\lambda < \dfrac{\delta}{||\hat{x}-x||}$ we have that

$$(1-\lambda)\bar{x} + \lambda\hat{x} \ \varepsilon\ B_\delta(\bar{x}) \cap P$$

and since we have a minimum at $\bar{x} \varepsilon B_\delta(\bar{x})$ it follows that

$$\theta(\bar{x}) \leqq \theta[(1-\lambda)\bar{x} + \lambda\hat{x}]$$

which contradicts (2.2.4) .

The above property of SQCX (SQCA) does not hold for QCX (QCA) functions, and is easily demonstrated by the numerical function θ defined on E^1 as follows:

$$\theta(x) = \begin{cases} x \text{ for } x \leqq 0 \\ 0 \text{ for } 0 < x < 1 \\ x-1 \text{ for } x \geqq 1 \end{cases}$$

θ is both QCX and QCA on E^1 and it is easy to see that $x = 1/2$ is both a local maximum and a local minimum, but not a global maximum or global minimum over E^1 .

Our next theorem gives a characterization of differentiable functions defined on an open convex set.

Theorem 2.2.5. Let θ be a numerical function defined on some open set $P \subset E^n$. Let P be convex and θ differentiable at $\bar{x} \varepsilon P$. If $\theta(\bar{x}) = \min_{x \varepsilon P} \theta(x)$

then

$$\nabla' \theta(\bar{x})(x-\bar{x}) \geqq 0$$

for all $x \varepsilon P$.

Proof: Let $x \varepsilon P$, and since P is convex we have for $0 \leqq \lambda \leqq 1$ $(1-\lambda)\bar{x} + \lambda x \varepsilon P$.

Since θ is differentiable at $\bar{x}$ and $\theta(\bar{x}) = \min_{x \varepsilon P} \theta(x)$

$$\theta(\bar{x}) \leqq \theta(x)$$

for all $x \varepsilon P$, and

$$0 \leqq \theta[(1-\lambda)\bar{x} + \lambda x] - \theta(\bar{x})$$

$$= \lambda \nabla' \theta(\bar{x})(x-\bar{x}) + \alpha[\bar{x}, \lambda(x-\bar{x})] \lambda ||x-\bar{x}||$$

where,

$$\lim_{\lambda \to 0} \alpha[\bar{x}, \lambda(x-\bar{x})] = 0 .$$

Hence, as $\lambda \to 0$ we have

$$\nabla' \theta(\bar{x})(x-\bar{x}) \geqq 0 .$$

Definition 2.2.7. A numerical function θ defined on an open set $P \subset E^n$ is said to be pseudoconvex (PCX) at $\bar{x} \varepsilon P$ if it is differentiable at $\bar{x}$ and

$$\left.\begin{array}{l} x \in P \\ \nabla'\theta(\bar{x})(x-\bar{x}) \geqq 0 \end{array}\right\rangle \text{implies} \quad \theta(x) \geqq \theta(\bar{x})$$

<u>Definition 2.2.8.</u> A numerical function θ defined on an open set $P \subset E^n$ is said to be pseudoconcave (PCA) at $\bar{x} \in P$ if it is differentiable at $\bar{x}$ and

$$\left.\begin{array}{l} x \in P \\ \nabla'\theta(\bar{x})(x-\bar{x}) \leqq 0 \end{array}\right\rangle \text{implies} \quad \theta(x) \leqq \theta(\bar{x})$$

θ is pseudoconcave at $\bar{x} \in P$ if and only if $-\theta$ is pseudoconvex at $\bar{x} \in P$.

<u>Theorem 2.2.6.</u> Let (i) P be convex in E^n and (ii) θ be a numerical function defined on an open set containing P. If θ is pseudoconvex on P, then θ is strictly quasiconvex on P.

<u>Proof:</u> Assume θ is PCX and that θ is not SQCX. This implies that there exists $x^1, x^2 \in P$ such that $\theta(x^2) < \theta(x^1)$ and $\theta(\hat{x}) \geqq \theta(x^1)$ for some $\hat{x} \in (x^1,x^2)$. Hence, there exists $\bar{x} \in (x^1,x^2)$ such that,

$$\theta(\bar{x}) = \max_{x \in [x^1,x^2]} \theta(x) \ .$$

By Theorem 2.2.5, we have

$$\nabla'\theta(\bar{x})(x^1-\bar{x}) \leqq 0 \tag{2.2.5}$$

and

$$\nabla'\theta(\bar{x})(x^2-\bar{x}) \leqq 0 \ . \tag{2.2.6}$$

Since $\bar{x} = (1-\lambda)x^1 + \lambda x^2$ for some $\lambda \in (0,1)$ we have in view of (2.2.5) that

$$0 \geqq \nabla'\theta(\bar{x})(x^1-\bar{x}) = \lambda\nabla'\theta(\bar{x})(x^1-x^2)$$

from (2.2.6)

$$0 \geqq \nabla'\theta(\bar{x})(x^2-\bar{x}) = -(1-\lambda)\nabla'\theta(\bar{x})(x^1 - x^2) .$$

Hence,

$$\nabla'\theta(\bar{x})(x^1-x^2) = 0$$

and

$$\nabla'\theta(\bar{x})(x^2-\bar{x}) = 0 .$$

But, since θ is PCX on P, it follows that

$$\theta(x^2) \geqq \theta(\bar{x}) \quad (\text{since } \nabla'\theta(\bar{x})(x^2-\bar{x}) = 0)$$

and hence,

$$\theta(x^1) > \theta(\bar{x}) \quad (\text{since } \theta(x^1) > \theta(x^2)) .$$

This last inequality contradicts the earlier statement that

$$\theta(\bar{x}) = \max_{x \in [x^1,x^2]} \theta(x) .$$

Therefore, θ is SQCX on P and by Theorem 2.2.3, is also QCX on P.

To see that the converse is not necessarily true, consider the example $\theta(x) = x^3$, $x \in E^1$, which is SQCX on E^1 but is not PCX on E^1 .

Theorem 2.2.7. Let θ be a numerical function defined on an open set $P \subset E^n$. Let $x \in P$ and let θ be differentiable at $\bar{x}$, then

(i) if $\theta(x) \geqq \theta(\bar{x})$ for all $x \in P$

$$\Rightarrow \nabla'\theta(\bar{x}) = 0 ,$$

(ii) if θ is pseudoconvex at $\bar{x}$, then

$$\nabla'\theta(\bar{x}) = 0 \Rightarrow \theta(x) \geqq \theta(\bar{x})$$

for all $x \in P$; i.e. $\theta(\bar{x}) = \min_{x \in P} \theta(x)$.

Proof of (i): By Theorem 2.2.5, $\theta(x) \geqq \theta(\bar{x})$ for all $x \in P$, implies

$$\nabla'\theta(\bar{x})(x-\bar{x}) \geqq 0$$

for all $x \in P$. Since P is open, we have

$$x = \bar{x} - \delta\nabla'\theta(\bar{x})$$

for $x \in P$ and some $\delta > 0$. Then $\nabla'\theta(\bar{x})(\bar{x}-\delta\nabla'\theta(\bar{x})-\bar{x}) \geqq 0$, implies

$$-\delta\nabla'\theta(\bar{x})\nabla'\theta(\bar{x}) \geqq 0 ,$$

$$[\nabla'\theta(\bar{x})]^2 \leqq 0 ,$$

$$\nabla'\theta(\bar{x}) = 0 .$$

Proof of (ii): If θ is PCX at $\bar{x}$ then

$$\nabla'\theta(\bar{x})(x-\bar{x}) \geqq 0 \Rightarrow \theta(x) - \theta(\bar{x}) \geqq 0$$

for all $x \in P$. From (i) above we get

$$\nabla'\theta(\bar{x}) = 0 ,$$

$$\nabla'\theta(\bar{x})(x-\bar{x}) = 0 ,$$

and $\theta(x) \geqq \theta(\bar{x})$ for all $x \in P$.

Let ϕ and ψ be differentiable functions defined on an open set $P \subset E^n$, and let $\psi \neq 0$ on P.

<u>Theorem 2.2.8.</u> If $\theta = \phi/\psi$, where ϕ is convex at $\bar{x}$ and $\psi > 0$ is linear on R^n, then θ is pseudoconvex at $\bar{x}$.

<u>Proof:</u> Assume $\nabla'\theta(x)(x-\bar{x}) \geq 0$ then

$$\nabla'\theta(\bar{x})(x-\bar{x}) = \psi(\bar{x})\nabla'\phi(\bar{x})(x-\bar{x}) - \phi(\bar{x})\nabla'\psi(\bar{x})(x-\bar{x}) \geq 0 . \qquad (2.2.7)$$

Since ϕ is convex at $\bar{x}$ we have

$$\nabla'\phi(\bar{x})(x-\bar{x}) \leq \phi(x) - \phi(\bar{x}),$$

also ψ linear at $\bar{x}$ gives us

$$\nabla'\psi(\bar{x})(x-\bar{x}) = \psi(x) - \psi(\bar{x}) .$$

Therefore, (2.2.7) can be rewritten as

$$\psi(\bar{x})[\phi(x) - \phi(\bar{x})] - \phi(\bar{x})\ [\psi(x) - \psi(\bar{x})] \geq 0$$

$$\psi(\bar{x})\phi(x) \geq \phi(\bar{x})\ \psi(x)$$

$$\frac{\phi(x)}{\psi(x)} \geq \frac{\phi(\bar{x})}{\psi(\bar{x})}$$

$$\theta(x) \geq \theta(\bar{x}) .$$

Hence, θ is PCX at $\bar{x} \in P$.

<u>Corollary 2.2.1.</u> If $\theta = \phi/\psi$ where ϕ is concave at $\bar{x}$ and $\psi < 0$ is linear on E^n, then θ is pseudoconcave at $\bar{x}$.

<u>Theorem 2.2.9.</u> Let θ be a numerical function defined on some set $P \subset E^n$, and let θ be defined as $\theta(x) = \dfrac{b'x + b_o}{d'x + d_o}$ where $b_o, d_o \in E^1$,

$b \in E^n$, $d \in E^n$, then θ is both pseudoconvex and pseudoconcave on each convex $P \subset E^n$ on which $d'x + d_o \neq 0$.

Proof: Let $P_1 \cup P_2 = P$ where $P_1 = \{x \mid d'x + d_o > 0\}$ and $P_2 = \{x \mid d'x + d_o < 0\}$. Since $b'x + b_o$ and $d'x + d_o$ are linear they are both convex and concave on P. From Theorem 2.2.8 and Corollary 2.2.1 if we let $\phi = b'x + b_o$, $\psi = d'x + d_o$, then for $x \in P_1$ θ is PCX on P_1 and for $x \in P_2$ we have θ is PCA on P_2.

Definition 2.2.9. A numerical function θ defined on a set $P \subset E^n$ is said to be convex (CX) at $\bar{x} \in P$ if

$$\left.\begin{array}{l} x \in P \\ 0 \leqq \lambda \leqq 1 \\ (1-\lambda)\bar{x} + \lambda x \in P \end{array}\right\rangle \text{implies } (1-\lambda)\theta(\bar{x}) + \lambda\theta(x) \geqq \theta[(1-\lambda)\bar{x}+\lambda x]$$

Definition 2.2.10. A numerical function θ defined on a set $P \subset E^n$ is said to be concave (CA) at $\bar{x} \in P$ if

$$\left.\begin{array}{l} x \in P \\ 0 \leqq \lambda \leqq 1 \\ (1-\lambda)\bar{x} + \lambda x \in P \end{array}\right\rangle \text{implies } (1-\lambda)\theta(\bar{x}) + \lambda\theta(x) \leqq \theta[(1-\lambda)\bar{x}+\lambda x]$$

Theorem 2.2.10. Let θ be a numerical function defined on an open set $P \subset E^n$ and let θ be differentiable at $\bar{x} \in P$. If θ is convex at $\bar{x} \in P$, then

$$\theta(x) \geq \theta(\bar{x}) + \nabla'\theta(\bar{x})(x-\bar{x})$$

for each $x \in P$.

Proof: Let θ be convex at $\bar{x}$. Since P is open, there exists an open ball $B_\delta(\bar{x})$, which is contained in P. Let $x \neq \bar{x}$ be in P, then for some u, such that $0 < u < 1$ and $u < \delta/||x-\bar{x}||$, we have that

$$\hat{x} = \bar{x} + u(x-\bar{x}) = (1-u)\bar{x} + ux \in B_\delta(\bar{x}) \subset P .$$

Since θ is convex at $\bar{x}$, it follows from Definition 2.2.9 and the convexity of $B_\delta(\bar{x})$, that for $0 < \lambda \leqq 1$

$$(1-\lambda)\theta(\bar{x}) + \lambda\theta(\hat{x}) \geqq \theta[(1-\lambda)\bar{x} + \lambda\hat{x}]$$

or

$$\theta(\hat{x}) - \theta(\bar{x}) \geqq \frac{\theta[\bar{x} + \lambda(\hat{x}-\bar{x})] - \theta(\bar{x})}{\lambda}$$

$$= \frac{\lambda\nabla'\theta(\bar{x})(\hat{x}-\bar{x}) + \alpha[\bar{x}, \lambda(\hat{x}-\bar{x})]\ \lambda||\hat{x}-\bar{x}||}{\lambda}$$

$$= \nabla'\theta(\bar{x})(\hat{x}-\bar{x}) + \alpha[\bar{x}, \lambda(\hat{x}-\bar{x})]\ ||\hat{x}-\bar{x}|| .$$

Since

$$\lim_{\lambda\to 0} \alpha[\bar{x}, \lambda(x-\bar{x})] = 0 ,$$

taking the limit of the previous expression as $\lambda \to 0$ gives

$$\theta(\hat{x}) - \theta(\bar{x}) \geqq \nabla'\theta(\bar{x})(\hat{x}-\bar{x}) . \tag{2.2.8}$$

Since θ is convex at $\bar{x}$, and $x \in P$, and since $\hat{x} = (1-u)\bar{x} + ux$, we have by Definition 2.2.9 that

$$u[\theta(x) - \theta(\bar{x})] \geqq \theta(\hat{x}) - \theta(\bar{x}) . \tag{2.2.9}$$

But since

$$\hat{x} - \bar{x} = u(x-\bar{x}) \tag{2.2.10}$$

and $u > 0$, (2.2.8)-(2.2.10) imply that

$$\theta(x) - \theta(\bar{x}) \geqq \nabla'\theta(\bar{x})(x-\bar{x}) \ .$$

The converse theorem is given by Mangasarian (1969, p. 84).

Theorem 2.2.11. Let θ be a numerical function defined on some open set $P \subset E^n$, let $\bar{x} \in P$, and let θ be differentiable at $\bar{x}$. If θ is convex (concave) at $\bar{x}$, then θ is pseudoconvex (pseudoconcave) at $\bar{x}$.

Proof: Let θ be convex at $\bar{x}$. By Theorem 2.2.10 we have that

$$\theta(x) - \theta(\bar{x}) \geq \nabla'\theta(\bar{x})(x-\bar{x})$$

for all $x \in P$, hence,

$$\left.\begin{array}{l} x \in P \\ \\ \nabla'\theta(x)(x-\bar{x}) \geqq 0 \end{array}\right\rangle \Rightarrow \theta(x) \geqq \theta(\bar{x})$$

and θ is pseudoconvex at $\bar{x}$.

The converse is not necessarily true, as can be seen from the example $\theta(x) = x + x^3$, $x \in E^1$. θ is not convex on E^1 because $\nabla^2\theta(x) < 0$ for $x < 0$. However, θ is pseudoconvex on E because

$$\nabla'\theta(x) = 1 + 3x^2 > 0 \text{ and}$$

$$\begin{aligned} \nabla'\theta(\bar{x})(x-\bar{x}) \geqq 0 &\Rightarrow x - \bar{x} \geqq 0 \\ &\Rightarrow x^3 \geqq \bar{x}^3 \\ &\Rightarrow \theta(x) - \theta(\bar{x}) = x + x^3 - \bar{x} - \bar{x}^3 \geqq 0 \ . \end{aligned}$$

<u>Theorem 2.2.12.</u> Let P be a convex set in E^n, and let θ be a numerical function defined on some open set containing P. If θ is pseudoconvex (pseudoconcave) on P, then each local minimum (maximum) of θ on P is a global minimum (maximum) of θ on P.

<u>Proof:</u> By Theorem 2.2.6 θ is SQCX (SQCA) on P. By Theorem 2.2.4 each local minimum (maximum) of θ on P is also a global minimum (maximum) on P.

<u>Definition 2.2.11.</u> A numerical function θ defined on a set $P \subset E^n$ is said to be strictly convex (SCX) at $\bar{x} \in P$ if

$$\left.\begin{array}{l} x \in P \\ x \neq \bar{x} \\ 0 < \lambda < 1 \\ (1-\lambda)\bar{x} + \lambda x \in P \end{array}\right\} \text{implies } (1-\lambda)\theta(\bar{x}) + \lambda\theta(x) > \theta[(1-\lambda)\bar{x} + \lambda x]$$

<u>Definition 2.2.12.</u> A numerical function θ defined on a set $P \subset E^n$ is said to be strictly concave (SCA) at $\bar{x} \in P$ if

$$\left.\begin{array}{l} x \in P \\ x \neq \bar{x} \\ 0 < \lambda < 1 \\ (1-\lambda)\bar{x} + \lambda x \in P \end{array}\right\} \text{implies } (1-\lambda)\theta(\bar{x}) + \lambda\theta(x) < \theta[(1-\lambda)\bar{x} + \lambda x]$$

A strictly convex (strictly concave) function on a set $P \subset E^n$ is obviously convex (concave) on P. The converse is not necessarily true, as can be seen from the example $\theta(x) = \alpha$, α a constant in E^1. θ is both convex and concave on E^n, but neither strictly convex nor strictly concave on E^n.

The properties and relationships between strictly convex, convex, pseudoconvex, strictly-quasiconvex, and quasiconvex functions are summarized in Diagram 2.2.1. The relationships hold under the assumption that θ is differentiable for each class of functions, θ is lower semicontinuous, and defined on an open convex set $P \subset E^n$.

If all the inequalities were reversed in Diagram 2.2.1 then the word concave could be substituted for the word convex.

Strictly Convex

$$(1-\lambda)\theta(\bar{x}) + \lambda\theta(x) > \theta[(1-\lambda)\bar{x} + \lambda x]$$

$$\theta(x) - \theta(\bar{x}) > \nabla'\theta(\bar{x})(x-\bar{x})$$

$$\Downarrow$$

Convex

$$(1-\lambda)\theta(\bar{x}) + \lambda\theta(x) \geqq \theta[(1-\lambda)\bar{x} + \lambda x]$$

$$\theta(x) - \theta(\bar{x}) \geqq \nabla'\theta(\bar{x})(x-\bar{x})$$

$$\Downarrow$$

Pseudoconvex

$$\nabla'\theta(\bar{x})(x-\bar{x}) \geqq 0 \Rightarrow \theta(x) \geqq \theta(\bar{x})$$

$$\Downarrow$$

Strictly Quasiconvex

$$\theta(x) < \theta(\bar{x}) \Rightarrow \theta[(1-\lambda)\bar{x} + \lambda x] < \theta(\bar{x})$$

$$\Downarrow$$

Quasiconvex

$$\theta(x) \leqq \theta(\bar{x}) \Rightarrow \theta[(1-\lambda)\bar{x} + \lambda x] \leqq \theta(\bar{x})$$

Diagram 2.2.1. Properties and Relationships of Convex Functions

2.3. Definitions of Generalized Convex Functions Over Cone Domains

In Section 2.2 we gave several definitions of generalized convex functions using inequalities. These definitions are now extended to arbitrary cones. C will denote a cone and "Int" will denote interior.

<u>Definition 2.3.1.</u> A set $C \subset E^m$ is a cone if $ky \in C$ for any $y \in C$, k (scalar) ≥ 0 .

<u>Definition 2.3.2.</u> A cone $C \subset E^m$ is convex if $(1-\lambda)y^1 + \lambda y^2 \in C$ for any two vectors y^1 and y^2 in C and $\lambda \in [0,1]$.

<u>Definition 2.3.3.</u> An m-dimensional vector function g defined on a set $P \subset E^n$ is said to be quasiconvex at $\bar{x}$ with respect to a cone C on P, if

$$\left.\begin{array}{l} x \in P \\ g(x) - g(\bar{x}) \in C \\ 0 \leqq \lambda \leqq 1 \\ (1-\lambda)\bar{x} + \lambda x \in P \end{array}\right\rangle \Rightarrow g[(1-\lambda)\bar{x} + \lambda x] - g(\bar{x}) \in C$$

<u>Definition 2.3.4.</u> An m-dimensional vector function g defined on a set $P \subset E^n$ is said to be quasiconcave at $\bar{x}$ with respect to a cone C on P, if

$$\left.\begin{array}{l} x \in P \\ g(\bar{x}) - g(x) \in C \\ 0 \leqq \lambda \leqq 1 \\ (1-\lambda)\bar{x} + \lambda x \in P \end{array}\right\rangle \text{implies } g(\bar{x}) - g[(1-\lambda)\bar{x} + \lambda x] \in C$$

g is said to be quasiconvex (quasiconcave) on P with respect to C if it is quasiconvex (quasiconcave) with respect to C at each $x \in P$.

Definition 2.3.5. An m-dimensional vector function g defined on a set $P \subset E^n$ is said to be strictly-quasiconvex at $\bar{x}$ with respect to a cone C on P, if

$$\left.\begin{array}{l} x \in P \\ g(x) - g(\bar{x}) \in \text{Int } C \\ 0 \leqq \lambda \leqq 1 \\ (1-\lambda)\bar{x} + \lambda x \in P \end{array}\right\rangle \text{implies } g[(1-\lambda)\bar{x} + \lambda x] - g(\bar{x}) \in \text{Int } C$$

Definition 2.3.6. An m-dimensional vector function g defined on a set $P \subset E^n$ is said to be strictly-quasiconcave at $\bar{x}$ with respect to a cone C on P, if

$$\left.\begin{array}{l} x \in P \\ g(\bar{x}) - g(x) \in \text{Int } C \\ 0 \leqq \lambda \leqq 1 \\ (1-\lambda)\bar{x} + \lambda x \in P \end{array}\right\rangle \text{implies } g(\bar{x}) - g[(1-\lambda)\bar{x} + \lambda x] \in \text{Int } C$$

Definition 2.3.7. Let g be an m-dimensional vector function defined on an open set $P \subset E^n$. g is pseudoconvex at $\bar{x}$ with respect to a cone C on P if g is differentiable at $\bar{x}$ and if

$$\left.\begin{array}{l} x \in P \\ \nabla g(\bar{x})(x-\bar{x}) \notin \text{Int } C \end{array}\right\rangle \text{implies } g(x) - g(\bar{x}) \notin \text{Int } C$$

<u>Definition 2.3.8.</u> Let g be an m-dimensional vector function defined on an open set $P \subset E^n$. g is pseudoconcave at $\bar{x}$ with respect to a cone C on P if g is differentiable at $\bar{x}$ and if

$$\left.\begin{array}{l} x \in P \\ \nabla g(\bar{x})(\bar{x}-x) \notin \text{Int } C \end{array}\right\rangle \text{ implies } g(\bar{x}) - g(x) \notin \text{Int } C$$

<u>Definition 2.3.9.</u> An m-dimensional vector function g defined on a set $P \subset E^n$ is said to be convex at $\bar{x} \in P$ with respect to a cone C on P if

$$\left.\begin{array}{l} x \in P \\ 0 \leqq \lambda \leqq 1 \\ (1-\lambda)\bar{x} + \lambda x \in P \end{array}\right\rangle \begin{array}{l} \text{implies } g[(1-\lambda)\bar{x} + \lambda x] \\ \qquad - [(1-\lambda)g(\bar{x}) + \lambda g(x)] \in C \end{array}$$

<u>Definition 2.3.10.</u> An m-dimensional vector function g defined on a set $P \subset E^n$ is said to be concave at $\bar{x} \in P$ with respect to a cone C on P if

$$\left.\begin{array}{l} x \in P \\ 0 \leqq \lambda \leqq 1 \\ (1-\lambda)\bar{x} + \lambda x \in P \end{array}\right\rangle \begin{array}{l} \text{implies } [(1-\lambda)g(\bar{x}) + \lambda g(x)] \\ \qquad - g[(1-\lambda)\bar{x} + \lambda x] \in C \end{array}$$

<u>Definition 2.3.11.</u> An m-dimensional vector function g defined on a set $P \subset E^n$ is said to be strictly convex at $\bar{x} \in P$ with respect to a cone C on P if

$$\left.\begin{array}{l} x \in P \\ x \neq \bar{x} \\ 0 < \lambda < 1 \\ (1-\lambda)\bar{x} + \lambda x \in P \end{array}\right\rangle \text{implies } g[(1-\lambda)\bar{x} + \lambda x] - [(1-\lambda)g(\bar{x}) + \lambda g(x)] \in \text{Int } C$$

<u>Definition 2.3.12.</u> An m-dimensional vector function g defined on a set $P \subset E^n$ is said to be strictly concave at $\bar{x} \in P$ with respect to a cone C on P if

$$\left.\begin{array}{l} x \in P \\ x \neq \bar{x} \\ 0 < \lambda < 1 \\ (1-\lambda)\bar{x} + \lambda x \in P \end{array}\right\rangle \text{implies } [(1-\lambda)g(\bar{x}) + \lambda g(x)] - g[(1-\lambda)\bar{x} + \lambda x] \in \text{Int } C$$

The following theorem, although presented in this section, is referred to in Section 2.4.

<u>Theorem 2.3.1.</u> If C is a convex cone and if y^1 and y^2 are in C, then the sum $y^1 + y^2$ is in C .

<u>Proof:</u> Since C is convex $y^1 + 1/2(y^2-y^1) = \hat{y} \in C$. By Definition 2.3.1, $2\hat{y} \in C$ and

$$2\hat{y} = y^1 + y^2 ,$$

therefore, $y^1 + y^2 \in C$.

2.4. Fritz John and Kuhn-Tucker Problems

Consider the following nonlinear programming problem:

(MP) Minimize $f(x)$

$$\text{subject to} \quad g(x) \leqq 0 \ , \ x \in P$$

where P is an open subset of E^n, and f and g are differentiable on P .

The best known optimality criteria for nonlinear programming problems are due to Fritz John, and Kuhn and Tucker. In case of necessary optimality criteria, the only restriction on the program (MP) is that g should satisfy a certain qualification. Several are given in Mangasarian (1969). For sufficient optimality criteria to hold, both f and g are required to satisfy certain convexity requirements. The Fritz John and the Kuhn-Tucker stationary point problems are defined below.

Fritz John stationary point problem:

(FJSP) Find $\bar{x} \in P$, $\bar{r}_o \in E^1$, and $\bar{r} \in E^m$, if they exist, such that

$$\bar{r}_o \nabla' f(\bar{x}) + \bar{r}' \nabla g(\bar{x}) = 0$$

$$g(\bar{x}) \leqq 0$$

$$\bar{r}' g(\bar{x}) = 0$$

$$\bar{r}_o > 0, \bar{r} \geqq 0 \ .$$

The equivalent Kuhn-Tucker stationary point problem is:

(KTSP) Find $\bar{x} \in P$, $\bar{r} \in E^m$, if they exist, such that

$$\nabla' f(\bar{x}) + \bar{r}'\nabla g(\bar{x}) = 0$$

$$g(\bar{x}) \leqq 0$$

$$\bar{r}'g(\bar{x}) = 0$$

$$\bar{r} \geqq 0 .$$

Sufficient optimality conditions for (MP) are given, by Mangasarian (1969).

<u>Theorem 2.4.1.</u> If f is PSX and g is QCX on P and if there exists $\bar{x} \varepsilon P$, $\bar{r}_o \varepsilon E^1$, and $\bar{r} \varepsilon E^m$ which satisfy the (FJSP) problem, then $\bar{x}$ solves (MP).

The equivalent Kuhn-Tucker condition would be:

<u>Theorem 2.4.2.</u> If f is PSX and g is QCX on P and if there exists $\bar{x} \varepsilon P$ and $\bar{r} \varepsilon E^m$ which satisfy the (KTSP) problem, then $\bar{x}$ solves (MP).

Now let us consider the following programming problem:

(MMP) Minimize $f(x)$

$$\text{subject to } g(x) \leqq 0$$

$$h(x) = 0$$

$$x \varepsilon P$$

where P is an open subset of E^n, and f, g, and h are differentiable on P.

Theorem 2.4.1 and 2.4.2 are applicable only when h is quasi-monotonic on P (i.e. QCX and QCA both). But these optimality criteria cannot be used for any other form of h.

Bhatt and Misra (1975) established sufficient optimality criteria of the type mentioned above in the case where f, g, and h are all convex functions. The following results are due to Bhatt and Misra.

<u>Theorem 2.4.3.</u> If f, g, and h are convex on P, and if there exists $\bar{x} \varepsilon P$, $\bar{r} \varepsilon E^m$, $\bar{s} \varepsilon E^k$ such that

$$\nabla' f(\bar{x}) + \bar{r}'\nabla g(\bar{x}) + \bar{s}'\nabla h(\bar{x}) = 0$$

$$g(\bar{x}) \leqq 0$$

$$h(\bar{x}) = 0$$

$$\bar{r}'g(\bar{x}) = 0$$

$$\bar{r} \geqq 0, \; \bar{s} \geqq 0$$

then $\bar{x}$ solves problem (MMP).

The equivalent modified Fritz John type sufficient condition would be:

<u>Theorem 2.4.4.</u> If f, g, and h are convex on P and if there exists $\bar{x} \varepsilon P$, $\bar{r}_o \varepsilon E^1$, $\bar{r} \varepsilon E^m$, $\bar{s} \varepsilon E^k$ such that

$$\bar{r}_o\nabla' f(\bar{x}) + \bar{r}'\nabla g(\bar{x}) + \bar{s}'\nabla h(\bar{x}) = 0$$

$$g(\bar{x}) \leqq 0$$

$$h(\bar{x}) = 0$$

$$\bar{r}'g(\bar{x}) = 0$$

$$\bar{r}_o > 0, \bar{r} \geqq 0, \; \bar{s} \geqq 0$$

then $\bar{x}$ solves problem (MMP).

Bhatt and Misra (1975) also considered the case when $\bar{x}$, $\bar{r}_o$, $\bar{r}$, $\bar{s}$ satisfy the condition of Theorem 2.4.4 except the requirement $\bar{r}_o > 0$. Not requiring $\bar{r}_o > 0$ but only that

$$(\bar{r}_o, \bar{r}, \bar{s}) \geq 0 \,, \quad (\bar{r}_o, \bar{r}, \bar{s}) \neq 0$$

and g and h to be strictly convex, they proved the following theorem.

<u>Theorem 2.4.5.</u> If f is convex and g and h are strictly convex on P and if there exists $\bar{x} \in P$, $\bar{r}_o \in E^1$, $\bar{r} \in E^m$, $\bar{s} \in E^k$ such that

$$\begin{aligned} \bar{r}_o \nabla' f(\bar{x}) + \bar{r}' \nabla g(\bar{x}) + \bar{s}' \nabla h(\bar{x}) &= 0 \\ g(\bar{x}) &\leqq 0 \\ h(\bar{x}) &= 0 \\ \bar{r}' g(\bar{x}) &= 0 \\ \bar{r}_o \geqq 0,\ \bar{r} \geqq 0,\ \bar{s} \geqq 0,\ (\bar{r}_o, \bar{r}, \bar{s}) &\neq 0 \end{aligned}$$

then $\bar{x}$ solves problem (MMP).

Skarpness and Sposito (1980) extend the results of Bhatt and Misra (1975) by considering f to be PCX with g and h defined as strictly pseudoconvex functions.

<u>Definition 2.4.1.</u> A numerical function θ defined on an open set $P \subset E^n$ which is differentiable at $\bar{x} \in P$ is said to be strictly pseudoconvex (SPCX) at $\bar{x}$ if

$$\left.\begin{array}{l} x \in P \\ x \neq \bar{x} \\ \theta(x) \leqq \theta(\bar{x}) \end{array}\right\} \text{ implies } \nabla\theta(\bar{x})(x-\bar{x}) < 0 \qquad (2.4.1)$$

This definition is a slight extension of a strictly convex function given in Section 2.2. Using this definition, we establish a sufficient optimality criteria of the Fritz John type.

<u>Theorem 2.4.6.</u> Let (i) f be pseudoconvex, g and h strictly pseudoconvex, and are all differentiable at $\bar{x} \in P$, and (ii) there exists

$$\bar{r}_o \in E^1,\ \bar{r} \in E^m,\ \bar{s} \in E^k$$

such that

$$\bar{r}_o \nabla' f(\bar{x}) + \bar{r}' \nabla g(\bar{x}) + \bar{s}' \nabla h(\bar{x}) = 0$$

$$g(\bar{x}) \leqq 0$$

$$h(\bar{x}) = 0$$

$$\bar{r}' g(\bar{x}) = 0$$

$$(r_o, \bar{r}, \bar{s}) \geq 0,\ (r_o, \bar{r}, \bar{s}) \neq 0$$

then $\bar{x}$ solves problem (MMP).

<u>Proof:</u> Let $I = \{i \mid g_i(\bar{x}) = 0\}$, $J = \{i \mid g_i(\bar{x}) < 0\}$,

$$I \cup J = \{1, 2, \ldots, m\} .$$

Since $\bar{r} \geqq 0$, $g(\bar{x}) \leqq 0$, and $\bar{r}' g(\bar{x}) = 0$ we have that $\bar{r}_i' g_i(\bar{x}) = 0$ for $i = 1, \ldots, m$ and hence, $\bar{r}_i = 0$ for $i \in J$.

Therefore, we can rewrite the first condition as

$$\bar{r}_o \nabla' f(\bar{x}) + \bar{r}_I' \nabla g_I(\bar{x}) + \bar{s}' \nabla h(\bar{x}) = 0 \ .$$

Appealing to Gordon's theorem [Theorem 2.4.5, page 31, Mangasarian (1969)] with

$$A = \begin{pmatrix} \nabla' f(\bar{x}) \\ \nabla g_I(\bar{x}) \\ \nabla h(\bar{x}) \end{pmatrix} \quad \text{and} \quad x = -z$$

there does not exist any $z \in P$ such that

$$\nabla' f(x) z < 0, \ \nabla g_I(\bar{x}) z < 0, \ \nabla h(\bar{x}) z < 0 \ . \tag{2.4.2}$$

Therefore, the system

$$f(x) - f(\bar{x}) < 0, \ g_I(x) - g_I(\bar{x}) \leqq 0, \ h(x) - h(\bar{x}) = 0 \tag{2.4.3}$$

has no solution $x \in P$. If there did exist a solution $x^o \in P$, $(x^o \neq \bar{x})$ then

$$f(x^o) - f(\bar{x}) < 0 \Rightarrow \nabla' f(\bar{x})(x^o - \bar{x}) < 0 \ \text{(PCX)}$$

$$g_I(x^o) - g_I(\bar{x}) \leqq 0 \Rightarrow \nabla g_I(\bar{x})(x^o - \bar{x}) < 0 \ \text{(from 2.4.1)}$$

$$h(x^o) - h(\bar{x}) = 0 \Rightarrow \nabla h(\bar{x})(x^o - \bar{x}) < 0 \ \text{(from 2.4.1)} \ .$$

But this violates (2.4.2) with $z = x^o - \bar{x}$. Hence, $\bar{x}$ is an optimal solution of problem (MMP), in view of (2.4.3) .

3. NECESSARY AND SUFFICIENT OPTIMALITY CONDITIONS

3.1. Introduction

Kuhn and Tucker (1951) derived a duality relationship between certain constrained optimization problems and related Lagrangian saddlepoint problems. Their results were established for problems over orthant domains. Specifying Lagrangian functions in this general context lead to the duality results.

Related work by Varaiya (1967) and Guignard (1969) generalized Kuhn-Tucker duality from finite dimensional orthant domains to domains in Banach spaces typically involving cones or local cones. Abrams (1973), Sposito (1974 and 1976) have established various optimality conditions for linear and quadratic programming problems replacing orthant domains by cone domains.

In Section 3.2 we will extend several results of Alders (1976) and Sposito (1976) by enlarging the class of objective functions to include pseudoconvex functions and constraints defined over both degenerate and nondegenerate cone domains. We will use the results in Section 3.3 to establish a modified Farkas Lemma over degenerate and nondegenerate cone domains. Furthermore, in Section 3.4, a quadratic programming problem over mixed cone domains is considered. Its dual problem is constructed in a natural way with degenerate and nondegenerate cone domains and strong duality results are established between the original problem, its dual, certain optimality conditions, and an associated saddle value problem.

3.2. Optimality Conditions

Let P be an open subset of E^n and C be an arbitrary cone in E^m.

Definition 3.2.1. C^* will denote the polar cone of an arbitrary cone C in E^m, that is,

$$C^* = \{y^* \varepsilon E^m : y^{*\prime}y \geq 0 \text{ for all } y \varepsilon C\} .$$

Definition 3.2.2. A cone C is pointed if $C \cap (-C) = \{0\}$.

Consider the following minimization problem:

Problem P: Find an $x^o \varepsilon E^n$, if it exists, such that

$$F(x^o) = \min_{x \varepsilon X} F(x) \ , \ x^o \varepsilon X$$

where

$$X = \{x : x \varepsilon P \subset E^n \ , \ g(x) \varepsilon C \subset E^m\} .$$

Associated with this minimization problem, Problem P, is a modified Kuhn-Tucker stationary point problem over cone domains, Mangasarian (1969). Find an $x^o \varepsilon P \subset E^n$ and $u^o \varepsilon -C^* \subset E^m$ such that

$$\nabla F(x^o) + u^{o\prime}\nabla g(x^o) = 0 \qquad (3.2.1)$$

$$u^{o\prime}g(x^o) = 0 \qquad (3.2.2)$$

$$g(x^o) \varepsilon C . \qquad (3.2.3)$$

Equations (3.2.1)-(3.2.3) will be denoted as the Modified Kuhn-Tucker conditions over cone domains.

In the remainder of this chapter, F will denote a pseudoconvex objective function and g will denote a system of linear equations

$c-A'x$, where A' is a $m \times n$ matrix of rank m $(m \leq n)$ defined over an arbitrary cone domain.

Mangasarian (1969) established necessary optimality conditions associated with Problem P by appealing to Gordan's Theorem of the Alternative. We will use a similar approach in this chapter but will appeal to a generalized Gordan's Theorem to establish similar necessary conditions. Berman and Ben-Israel (1971) generalized Gordan's Theorem to problems defined over pointed, closed convex cone domains using a special case of Mazur's theorem, see Bourbaki (1953), p. 69. We will establish a generalized Gordan's theorem without relying on Mazur's theorem.

Lemma 3.2.1. Let R be a closed cone with polar R^*. If the interior of R^* is nonempty, then R is a pointed cone.

Proof: Assume R is not pointed then for some $y \in R$ we have that $-y \in R$. Moreover, for some $y^* \in \text{Int}(R^*)$ we have that $y^{*\prime}y > 0$ and $y^{*\prime}(-y) > 0$. A contradiction and it follows that if $\text{Int}(R^*) \neq \phi$, then R is pointed.

Lemma 3.2.2. If a cone R is convex, pointed, and closed, then $\text{Int}(R^*) \neq \phi$.

Proof: Assume $\text{Int}(R^*) = \phi$, then for $y^* \in R^*$ there exists $\bar{y} \in R$, such that $y^{*\prime}\bar{y} = 0$ and $-(y^{*\prime}\bar{y}) = 0$. Since R is pointed, then necessarily $-y^* \in R^*$, and R^* is not pointed. Therefore, we have from Lemma 3.2.1 that $\text{Int}(R) = \phi$.

Now $\mathrm{Int}(R) = \phi$ and $\mathrm{Int}(R^*) = \phi$ imply that $R = \{y \mid y'y^* = 0 \ \forall\ y^* \in R^*\}$. Hence, if $\bar{y} \in R$, then $-\bar{y} \in R$ and we have that R is not pointed. A contradiction and the result follow.

Theorem 3.2.1. (Gordan's Theorem for cone domains). Let M be any given nonvacuous $m \times n$ matrix with R any arbitrary cone in E^n which is closed, convex, and pointed, then exactly one of the following systems is consistent.

(a) $Mx = 0$ for some $x \in R$, $x \neq 0$

or

(b) $M'y \in \mathrm{Int}(-R^*)$, $y \in E^m$.

Proof: (Not (b) implies (a)). Let $S_1 = \{z \mid z = M'y,\ y \in E^m\}$ and $S_2 = \{z \mid z \in \mathrm{Int}(-R^*)\}$, then $S_1 \cap S_2 = \phi$; S_1 and S_2 are convex sets. Therefore, there exists a hyperplane v (nonzero), such that

$$v'z_1 \geq v'z_2 \ \forall\ z_1 \in S_1 \ ; \ \forall\ z_2 \in \bar{S}_2, \text{ (the closure of } S_2).$$

Hence, $v'M'y \geq v'z_2 \ \forall\ y \in E^m$; $\forall\ z_2 \in \bar{S}_2$. Assume $v \notin R$, then there exists $z_2^* \in \bar{S}_2$ such that $v'z_2^* > 0$. Moreover, for any given $y^* \in E^m$, there exists a $\bar{z}_2^* = kz_2^* \in \bar{S}_2$, $k > 0$, such that $v'\bar{z}_2^* > v'M'y^*$. Hence, it follows that $v \in R$.

Now letting $\bar{z} = 0$, then $v'M'y \geq v'\bar{z} = 0$, hence, $v'M'y \geq 0$. However, letting $y = -Mv$, then $-v'M'Mv \geq 0$ or $Mv = 0$. Therefore, $Mv = 0$, $v \in R$ $(v \neq 0)$, hence, (a) holds.

((b) implies not (a)). Let y^* be such that $M'y^* \in \mathrm{Int}(-R^*)$ and assume there exists $x^* \in R$ such that $Mx^* = 0$; $x^* \neq 0$, then

$y^{*\prime}(Mx^{*}) = 0$. A contradiction, since $M'y^{*} \in \mathrm{Int}(-R^{*})$ with $x^{*} \in R$, $(x^{*} \neq 0)$, implies that $x^{*\prime}(M'y^{*}) < 0$, hence, the result follows.

We now define a condition which is necessary to establish subsequent results.

Condition I: C is a closed, convex cone with nonempty interior.

Theorem 3.2.2. (Kuhn-Tucker Necessary Conditions). Assume C satisfies Condition I in Problem P. If x^{o} is an optimal solution of Problem P, then there exists u^{o}, such that (x^{o},u^{o}) satisfies conditions (3.2.1)-(3.2.3).

Proof: Let x^{o} solve Problem P, and assume there is no $(\bar{r}_{o},\bar{r}) \in Q_1^{+} \times (-C^{*})$; $(\bar{r}_{0},\bar{r}) \neq 0$ such that

$$\bar{r}_{o}\nabla' F(x^{o}) - \bar{r}'\nabla g(x^{o}) = 0$$

$$\bar{r}' g(x^{o}) = 0 \quad \text{where} \quad g(x) = c - A'x \,.$$

Now $\mathrm{Int}(C) \neq \phi$, therefore in view of Lemma 3.2.1 with $C = R^{*}$ and $C^{*} = R$, we have that C^{*} is pointed. Moreover, the cone $Q_1^{+} \times (-C^{*})$ is pointed, therefore, from Gordan's Theorem, (b) holds. In particular, letting

$$M = \begin{pmatrix} \nabla F(x^{o}) & \nabla' g(x^{o}) \\ 0 & g'(x^{o}) \end{pmatrix}; \; x = \begin{pmatrix} \bar{r}_{o} \\ \bar{r} \end{pmatrix} \in \begin{pmatrix} Q_1^{+} \\ -C^{*} \end{pmatrix} = R \,,$$

then there exists $y = (y_1, y_2)$; $y_1 \in E^{n}$, $y_2 \in E^{1}$ such that

$$\left.\begin{aligned} &\nabla' F(x^{o})y_1 \in \mathrm{Int}\, Q_1^{-} \\ &\nabla g(x^{o})y_1 + g(x^{o})y_2 \in \mathrm{Int}\, C \end{aligned}\right\} \qquad (3.2.4)$$

Now for sufficiently small $t > 0$ we have for y_1 and y_2 in (3.2.4) that

$$g(x^o+ty_1) - g(x^o) = t\nabla g(x^o)y_1 + \alpha(x^o,ty_1)||ty_1|| ,$$

and adding and subtracting $tg(x^o)y_2$ we obtain

$$g(x^o+ty_1) = t\nabla g(x^o)y_1 + tg(x^o)y_2 - tg(x^o)y_2 + g(x^o) + o(t)$$

$$= (1-ty_2)g(x^o) + t[\nabla g(x^o)y_1 + g(x^o)y_2] + o(t) .$$

Choosing $t > 0$ such that $1 - ty_2 \geq 0$, then $(1-ty_2)g(x^o) \in C$ and in view of (3.2.4)

$$t[\nabla g(x^o)y_1 + g(x^o)y_2] \in \text{Int } C ,$$

therefore,

$$g(x^o+ty_1) \in C .$$

Also,

$$F(x^o+ty_1) - F(x^o) = t\nabla' F(x^o)y_1 + o(t)$$

and with

$$\nabla' F(x^o)y_1 \in \text{Int } Q_1^- \text{ in (3.2.4)} ,$$

we have $F(x^o+ty_1) - F(x^o) \in \text{Int } Q_1^-$.

In summary,

$$F(x^o+ty_1) < F(x^o)$$

and $g(x^o+ty_1) \in C$ which implies that x^o is not optimal.

Therefore, there exists $(\bar{r}_o, \bar{r}) \in Q_1^+ \times -C^*$ such that

$$\bar{r}_o \nabla' F(x^o) + \bar{r}' \nabla g(x^o) = 0 \,, \tag{3.2.5}$$

$$\bar{r}' g(x^o) = 0 \,.$$

To establish now the existence of $u^o \in -C^*$ such that (x^o, u^o) satisfies the Kuhn-Tucker conditions. Assume $\bar{r}_o = 0$. Since $g(x)$ is linear with rank m ($m \leq n$) then clearly, there exists $y \in E^n$ such that

$$\nabla g(x^o) y \in \text{Int } (-C) \,.$$

Now with $0 \neq \bar{r} \in -C^*$, we have $\bar{r}'(\nabla g(x^o) y) > 0$.

However, $\bar{r}' \nabla g(x^o) = 0$ in view of (3.2.5) with $\bar{r}_o = 0$, a contradiction and $\bar{r}_o > 0$. Dividing (3.2.5) by $\bar{r}_o$, the result follows letting

$$u^o = \frac{\bar{r}}{\bar{r}_o} \in -C^* \,.$$

Sufficiency of the Kuhn-Tucker conditions is established in the following theorem.

Theorem 3.2.3. (Sufficiency of the Kuhn-Tucker conditions). If (x^o, u^o) satisfy the Kuhn-Tucker conditions, constructed from Problem P where C is an arbitrary cone, then x^o solves Problem P.

Proof: Let $g(x^o) = c - A'x^o \in C$. If (x^o, u^o) satisfies the Modified Kuhn-Tucker conditions, then

$$\nabla' F(x^o) + u^{o\prime} \nabla g(x^o) = 0$$

with $u^o \in -C^*$ or for any $x \in E^n$

$$\left[\nabla' F(x^o) + u^{o\prime} \nabla g(x^o) \right] (x - x^o) = 0 \,. \tag{3.2.6}$$

Moreover, since g is linear

$$\nabla g(x^o)(x-x^o) = [g(x) - g(x^o)]$$

for all $x \varepsilon E^m$ and since $u^{o\prime}g(x^o) = 0$,

then

$$u^{o\prime}\nabla g(x^o)(x-x^o) = u^{o\prime}g(x) \leq 0$$

for all feasible x .

Hence,

$$u^{o\prime}\nabla g(x^o)(x-x^o) \leq 0$$

implies that in (3.2.6)

$$\nabla' F(x^o)(x-x^o) \leq 0$$

for all feasible x , but since F is pseudoconvex,

it follows that

$$F(x^o) \leq F(x)$$

for all feasible x ; i.e. x^o solves Problem P.

3.3. Modified Farkas Lemma

In Section 3.2, optimality conditions associated with Problem P were derived. In this section, a modified Farkas Lemma over arbitrary cone domains is obtained appealing to these conditions by requiring only a "partial" linear dual theorem. These results are similar to those obtained by Sposito and David (1972) where the cone domains were non-degenerate.

Let $F(x) = b'x$ in Problem P, if x^o solves Problem P, then necessarily, in view of Theorem 3.2.2, there exists $u^o \varepsilon -C^*$ such that

$$b - Au^o = 0$$

$$u^{o\prime}(c - A'x^o) = 0$$

$$c - A'x^o \in C\,.$$

Consider now the following modified Farkas Lemma.

Lemma 3.3.1. Assume C satisfies Condition I. Then a vector $b \in E^n$ will satisfy $b'x \geq 0$ for all $x \in P$ with $A'x \in -C$ if there exists $u^o \in -C^*$ with $Au^o = b$.

Proof: Assume there exists $\bar{u} \in -C^*$ such that $A\bar{u} = b$, then for all $x \in E^n$ we have that $b'x = \bar{u}'A'x$. If in addition there exists $\bar{x}$ such that $A'\bar{x} \in -C$, then $\bar{u}'A'\bar{x} \geq 0$; hence, $b'\bar{x} \geq 0$.

Conversely, if $\{x \in E^n : A'x \in -C\}$ implies $b'x \geq 0$, then $x^o = 0$ solves Problem P with $c = 0$. Hence, under the assumption that rank $(A') = m$, there exists $u^o \in -C^*$ which satisfies (3.3.1)-(3.3.3); hence, $Au^o = b$.

The above lemma was established by Ben-Israel (1969b) using a different argument, [Theorem 2.4]. In particular, under the condition that the null space of $A + (-C^*)$ be closed.

A complete duality theory can be obtained utilizing certain results from Mangasarian (1969) and Sposito (1974) in addition to those in Section 3.2.

3.4. Quadratic Programming Problem

A "complete" quadratic duality theory for dual problems over degenerate and nondegenerate cone domains is now established. The problems are a special case of the duality theorem of complex quadratic programming [Theorem 4.1] considered by Abrams and Ben-Israel (1969), but generalized to constraints involving arbitrary cones. Our approach deviates from the quadratic duality results presented by Sposito (1976) where all the polar domains associated with each problem were required to be nonempty.

Consider the quadratic problem:

Problem QP: Maximize $G(x,u) = -x'Dx + c'u$

$$\text{subject to } h(x,u) = 2Dx - Au + b = 0$$

$$u \varepsilon -C^*$$

$$x \varepsilon E^n$$

where D is an $n \times n$ symmetric positive semi-definite matrix.

The proposed dual problem is:

Problem QD: Minimize $F(x) = x'Dx + b'x$

$$\text{subject to } c - A'x \varepsilon C$$

$$x \varepsilon E^n$$

where D is an $n \times n$ symmetric positive semi-definite matrix.

Our first objective is to show that if x^o solves Problem QD then there exists $u^o \varepsilon -C^*$ such that (x^o,u^o) solves Problem QP.

Assume that in Problem QD, that C is a closed convex cone with nonempty interior. If x^o is an optimal solution of Problem QD, then by Theorem 3.2.2 there exists u^o, such that (x^o,u^o) satisfies the modified Kuhn-Tucker conditions. For Problem QD these conditions would be:

$$2Dx^o - Au^o + b = 0 \tag{3.4.1}$$

$$u^{o\prime}(c-A'x^o) = 0 \tag{3.4.2}$$

$$c' - A'x^o \in C \tag{3.4.3}$$

$$u^o \in -C^*. \tag{3.4.4}$$

From (3.4.1)-(3.4.4) we have immediately that (x^o,u^o) is a feasible solution of Problem QP. Let us now establish some preliminary results which will be used to prove our first objective.

<u>Lemma 3.4.1.</u> (Weak Duality). If $\bar{x}$ is a feasible solution of Problem QD and $(\hat{x},\hat{u}) \in E^n \times -C^*$ such that $(\hat{x},\hat{u})$ satisfies (3.4.1) then

$$\bar{x}'D\bar{x} + b'\bar{x} \geq - \hat{x}'D\hat{x} + c'\hat{u} .$$

<u>Proof:</u> Since $(\hat{x},\hat{u})$ satisfies $2D\hat{x} - A\hat{u} + b = 0$ we have that $2\bar{x}'D\hat{x} - \bar{x}'A\hat{u} + \bar{x}'b = 0$. With $\hat{u} \in -C^*$ and $c - A'\bar{x} \in C$ we have

$$c'\hat{u} - \bar{x}'A\hat{u} \leq 0$$

which implies

$$2\bar{x}'D\hat{x} - \bar{x}'A\hat{u} + b'\bar{x} \geq c'\hat{u} - \bar{x}'A\hat{u} . \tag{3.4.5}$$

Since D is a symmetric positive semi-definite matrix,

$$(\bar{x}-\hat{x})'D(\bar{x}-\hat{x}) \geq 0$$

or

$$\bar{x}'D\bar{x} + \hat{x}'D\hat{x} \geq 2\bar{x}'D\hat{x} .$$

This implies from (3.4.5) that

$$\bar{x}'D\bar{x} + \hat{x}'D\hat{x} + b'\bar{x} \geq c'\hat{u}$$

therefore,

$$\bar{x}'D\bar{x} + b'\bar{x} \geq - \hat{x}'D\hat{x} + c'\hat{u} \ .$$

Corollary 3.4.1. If $\bar{x}$ and $(\tilde{x},\tilde{u})$ are feasible solutions of Problem QD and QP, respectively, and

$$\bar{x}'D\bar{x} + b'\bar{x} = - \tilde{x}'D\tilde{x} + c'\tilde{u}$$

then $\bar{x}$ is an optimal solution of Problem QD and $(\tilde{x},\tilde{u})$ is an optimal solution of Problem QP.

Proof: For any feasible solution x of Problem QD, we have from the Weak Duality lemma that

$$x'Dx + b'x \geq c'\tilde{u} - \tilde{x}'D\tilde{x} = \bar{x}'D\bar{x} + b'\bar{x}$$

thus, $x'Dx + b'x \geq \bar{x}'D\bar{x} + b'\bar{x}$ for any feasible solution x, that is $\bar{x}$ is an optimal solution of Problem QD.

Similarly, $(\tilde{x},\tilde{u})$ solves Problem QP.

Theorem 3.4.1. Under Condition I and the condition that the rank $(A') = m$, if x^o solves Problem QD then there exists u^o such that (x^o,u^o) solves Problem QP.

Proof: If x^o solves Problem QD then by Theorem 3.2.2, under the condition that rank $(A') = m$, there exists u^o, such that (x^o,u^o) satisfies the modified Kuhn-Tucker conditions. From (3.4.1) and (3.4.2) we have

$$2x^{o}Dx^{o} - x^{o\prime}Au^{o} + x^{o\prime}b = 0$$

and

$$u^{o\prime}c - u^{o\prime}A'x^{o} = 0 \ .$$

Hence,

$$x^{o\prime}Dx^{o} + b'x^{o} = - x^{o\prime}Dx^{o} + c'u^{o}$$

which implies from Corollary 3.4.1 that (x^{o},u^{o}) solves Problem QP.

The converse of the above theorem can be established appealing to and paralleling the arguments given by Sposito (1974) and Sposito (1976). In this vein consider and define the two sets K and V as

$$K = \left\{ \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} \varepsilon E^{n+1} \;\middle|\; \begin{array}{l} - t_1 - G(x,u) \leq 0 \\ h(x,u) - t_2 = 0 \end{array} \quad \text{for some } x \ \varepsilon\ E^{n},\ u\ \varepsilon\ -C^{*} \right\}$$

and

$$V = \left\{ \begin{pmatrix} t_1 \\ t_2 \end{pmatrix} \varepsilon E^{n+1} \;\middle|\; \begin{array}{r} - G(x^{o},u^{o}) - t_1 > 0 \\ t_2 = 0 \ \varepsilon\ E^{n} \end{array} \right\}$$

where (x^{o},u^{o}) is an optimal solution to Problem QP.

We now show that under certain conditions K is convex.

<u>Lemma 3.4.2.</u> If $G(x,u)$ is a concave function, $h(x,u)$ is a componentwise linear function of (x,u) and C is an arbitrary convex cone, then K is convex.

<u>Proof:</u> Let z^1 and $z^2 \in K$, then there exists $(x^1,u^1) = y^1 \in E^n \times -C^*$ such that

$$- z_1^1 - G(y^1) \leq 0 \tag{3.4.6}$$

and $(x^2,u^2) = y^2 \in E^n \times -C^*$ such that

$$- z_1^2 - G(y^2) \leq 0 . \tag{3.4.7}$$

multiplying (3.4.6) by α and (3.4.7) by $(1-\alpha)$ we have

$$- \tilde{z}_1 - \alpha G(y^1) - (1-\alpha)G(y^2) \leq 0 \tag{3.4.8}$$

for any $\alpha \in [0,1]$ where

$$\tilde{z}_1 = \alpha z_1^1 + (1-\alpha)z_1^2 .$$

But G is concave, therefore,

$$G(\alpha y^1 + (1-\alpha)y^2) \geq \alpha G(y^1) + (1-\alpha)G(y^2)$$

Hence, from (3.4.8)

$$- \tilde{z}_1 - G(\alpha y^1 + (1-\alpha)y^2) \leq 0$$

where $\alpha y^1 + (1-\alpha)y^2 \in E^n \times -C^*$. Also

$$h_i(y^1) - z_2^1 = 0 \quad \text{and} \quad h_i(y^2) - z_2^2 = 0, \; i = 1, \ldots, n .$$

But each h_i is linear, and it follows that

$$h_i(\alpha y^1 + (1-\alpha)y^2) - \hat{z}_2 = 0, \; i = 1, \ldots, n$$

where $\alpha y^1 + (1-\alpha)y^2 \in E^n \times -C^*$,

and K is convex.

Lemma 3.4.3. V is convex.

Proof: Let z^1 and z^2 be two arbitrary points in V, and let $(x^o,u^o) = y^o$ be an optimal solution of Problem QP, then

$$- G(y^o) - z_1^1 > 0 ,$$

$$z_2^1 = 0$$

and

$$- G(y^o) - z_1^2 > 0 ,$$

$$z_2^2 = 0 .$$

Consider, $\alpha z^1 + (1-\alpha)z^2 = \hat{z}$ for any $\alpha \varepsilon [0,1]$

then

$$\alpha[- G(y^o) - z_1^1] + (1-\alpha)[- G(y^o) - z_1^2] > 0 ,$$

$$-G(y^o) - (\alpha z_1^1 + (1-\alpha)z_1^2) > 0$$

that is

$$- G(y^o) - \hat{z}_1 > 0 .$$

Also,

$$\alpha z_2^1 + (1-\alpha)z_2^2 = 0$$

implies

$$\hat{z}_2 = 0$$

therefore V is convex, since $\hat{z}_1 \varepsilon E^1$ and $\hat{z}_2 \varepsilon E^n$.

<u>Lemma 3.4.4.</u> $K \cap Q = \phi$

<u>Proof:</u> Suppose $K \cap Q \neq \phi$; then there exists t^* such that $t^* \varepsilon K$ and $t^* \varepsilon V$. If $t^* \varepsilon K$, then for some $(x^1,u^1) \varepsilon E^n \times -C^*$ we have

$$- t_1^* - G(x^1,u^1) \leq 0$$

and

$$h(x^1,u^1) - t_2^* = 0 \ ,$$

If $t^* \varepsilon Q$, then $- G(x^o,u^o) - t_1^* > 0$, $t_2^* = 0$. It follows that $h(x^1,u^1) = 0$ and

$$G(x^o,u^o) < - t_1^* \leq G(x^1,u^1)$$

which contradicts the assumption that (x^o,u^o) is an optimal solution of the defined maximization problem.

<u>Lemma 3.4.5.</u> If $s'(d-Bx) \geq 0$ for some $s \neq 0$ and all $x \varepsilon E^n$, then $B's = 0$.

<u>Proof:</u> Assume that $B's \neq 0$. If $s'd \leq 0$, then pick $x^* = \delta B's/s'BB's$, where $\delta = \varepsilon + s'd$ for some $\varepsilon > 0$. This implies that $s'(d-Bx^*) < 0$. If $s'd < 0$, then for $x^* = 2B'ss'd/s'BB's$ we have $s'(d-Bx^*) < 0$. Hence, $B's = 0$.

Let $\Phi(x,u,z) = c'u - x'Dx + z'(2Dx-Au+b)$ be the Lagrangian function associated with Problem QP. Let us consider the following saddle value problem.

<u>Saddle Value Problem</u>: Find an $x^o \varepsilon E^n$, $u^o \varepsilon -C^*$, and $z^o \varepsilon E^n$. such that

$$\Phi(x,u,z^o) \leq \Phi(x^o,u^o,z^o) \leq \Phi(x^o,u^o,z)$$

for all $x \varepsilon E^n$, $u \varepsilon -C^*$, and $z \varepsilon E^n$.

Let us assume that the rank $h(x,u) = n$, then we can prove the following converse duality theorem.

<u>Theorem 3.4.2.</u> If (x^o,u^o) is a solution of Problem QP then there exists $z^o \varepsilon E^n$ such that

(a) (x^o,u^o,z^o) solves the saddle value problem,

(b) z^o solves Problem QD,

and moreover,

(c) $\Phi(x^o,u^o,z^o) = F(z^o)$.

<u>Proof</u>: (i) Since K and V are convex sets whose intersection is the null set, this implies there exists a separating hyperplane $v'z = \beta$, $v \neq 0$, such that

$$v'\hat{z} \geq v'\tilde{z} \tag{3.4.9}$$

where $\hat{z} \varepsilon K$ and $\tilde{z} \varepsilon \bar{V}$ ($\bar{V}$ is the closure of V relative to E^{n+1}), and

$$V = (v_1,v_2) \varepsilon L^* = \{v \mid v_1 \geq 0, v_2 \varepsilon E^n\}$$

(ii) We next establish that the first component of $v(v_1)$ is greater than or equal to zero. From (3.4.9), for any $(\hat{z},\tilde{z}) \varepsilon K \times \bar{V}$,

$$v_1\hat{z}_1 + v_2\hat{z}_2 \geq v_1\tilde{z}_1 + v_2\tilde{z}_2 = v_1\tilde{z}_1 \tag{3.4.10}$$

since $\tilde{z}_2 = 0$.

This implies if we add $v_1 G(x^o,u^o)$ to both sides of (3.4.10) that,

$$v_1(\hat{z}_1 + G(x^o,u^o)) + v_2\hat{z}_2 \geq v_1\tilde{z}_1 + v_1 G(x^o,u^o)$$

or

$$v_1\hat{z}_1 + v_2\hat{z}_2 + v_1 G(x^o,u^o) \geq v_1(\tilde{z} + G(x^o,u^o)) . \qquad (3.4.11)$$

Since $(\tilde{z} + G(x^o,u^o)) < 0$, if $v_1 \not\geq 0$ then it is possible to violate (3.4.11), therefore, $v \in L^*$.

(iii) To next establish that $v_1 > 0$, or in view of (ii) that $v_1 \neq 0$, we consider any $(x,u) \in E^n \times -C^*$; then $(G(x,u),h(x,u)) \in K$. Since $(G(x^o,u^o),0) \in \bar{V}$, (3.4.10) yields

$$v_1 G(x,u) + v_2 h(x,u) \geq v_1 G(x^o,u^o)$$

for all $(x,u) \in E^n \times -C^*$.

Now suppose that $v_1 = 0$; then we have

$$v_2' h(x,u) \geq 0$$

for all $(x,u) \in E^n \times -C^*$,

or

$$v_2'(2Dx-Au+b) \geq 0 \ .$$

If we let $B = [-2D : A]$, then from Lemma 3.4.5, $B'v_2 = 0$ for some $v_2 \neq 0$, contradicting the assumption that the rows of B are linear independent, that is the rank of $h(x,u) = n$. Thus, $v_1 \neq 0$.

Let $z^o = \frac{v_2}{v_1} \varepsilon E^n$, then we have,

$$G(x,u) + z^{o\prime}h(x,u) \geq G(x^o,u^o) \tag{3.4.12}$$

for all $(x,u) \varepsilon E^n \times -C^*$.

It remains to be shown that z^o is as specified in the statement of the theorem. Since $h(x^o,u^o) = 0$, and in view of (3.4.12)

$$\Phi(x,u,z^o) = G(x,u) + z^{o\prime}h(x,u) \geq G(x^o,u^o) = \Phi(x^o,u^o,z^o)$$

for all $(x,u) \varepsilon E^n \times -C^*$.

Also,

$$\Phi(x^o,u^o,z) = G(x^o,u^o) + z'h(x^o,u^o) \leq G(x^o,u^o) = \Phi(x^o,u^o,z^o)$$

for all $z \varepsilon E^n$, and we have established conclusion (a).

Now from Theorem 1 in Sposito (1976), the Kuhn-Tucker conditions necessarily hold; hence, from the sufficiency theorem, Theorem 3.2.3, it follows that z^o solves Problem QD.

We also have from the Kuhn-Tucker conditions of Theorem 1 in Sposito (1976) that

$$\nabla_x \Phi(x^o,u^o,z^o) = -2Dx^o + 2z^{o\prime}D = 0 .$$

Since $z^o \varepsilon E^n$, this implies

$$Dx^o = Dz^o . \tag{3.4.13}$$

Using (3.4.13) and (3.4.2) we have

$$\Phi(x^{o},u^{o},z^{o}) = - x^{o\prime}Dx^{o} + c'u^{o} + z^{o\prime}(2Dx^{o} - Au^{o} + b)$$

$$= - z^{o\prime}Dx^{o} + c'u^{o} + 2z^{o\prime}Dx^{o} - z^{o\prime}Au^{o} + z^{o\prime}b$$

$$= z^{o\prime}Dx^{o} + c'u^{o} - z^{o\prime}Au^{o} + b'z^{o}$$

$$= z^{o}Dx^{o} + b'z^{o}$$

$$= F(x^{o}) .$$

Hence, $\Phi(x^{o},u^{o},z^{o}) = F(x^{o})$ establishing part (c) .

4. LINEAR FRACTIONAL PROGRAMMING OVER CONE DOMAINS

4.1. Introduction

One of the first investigations of the computational and duality concepts of linear fractional programming was that of Charnes and Cooper (1962). Since then, the subject has been extensively researched. In this chapter we consider a linear fractional programming problem (LFP) in which the objective function, being the ratio of two appropriately restricted linear functions, is a pseudoconvex function and the constraints are linear inequalities defined over some arbitrary convex cone domain. A quadratic fractional program (QFP) is introduced, and used as a vehicle, along with a Weak Duality lemma and a Direct Duality theorem, to relate the (LFP) problem to a linear dual program (LDP). In particular, it is shown that the global optimum of the (LFP) can be obtained by solving the (LDP).

Bector (1973a) has similar concepts as those which appear in this chapter; in particular, with problems over orthant domains. Our results extends Bector's results by considering these problems over nondegenerate and degenerate arbitrary convex cone domains. In this vein, we shall appeal to the results developed in Chapter 3.

4.2. Linear Fractional Duality Formulation and Results

Consider the following linear fractional program (LFP), which we shall denote as the primal problem:

$$\text{(LFP):} \qquad \text{Minimize} \quad f(x) = \frac{b'x + b_o}{d'x + d_o}$$

$$\text{subject to} \quad c - A'x \in C$$

$$x \in E^n$$

where

(i) b_o, $d_o \in E$ are known constants,

(ii) b, d $\in E^n$ are known constants,

(iii) A' is an $m \times n$ matrix of rank m ,

(iv) $d'x + d_o > 0$ for all feasible x ,

(v) C is a nonempty convex cone, subset of E^m ,

(vi) $X = \{x : c - A'x \in C, x \in E^n\}$ is nonempty and bounded.

Associated with (LFP) problem is a quadratic fractional program (QFP), which we introduce as:

$$\text{(QFP):} \qquad \text{Maximize} \quad F(x,v) = \frac{b'x + b_o}{d'x + d_o} + v'(c-A'x)$$

$$\text{subject to} \quad (x,v) \in D$$

where the set D is nonempty and given by,

$$D = \{(x,v) : \nabla_x F(x,v) = 0, x \in E^n, v \in -C^*\} .$$

We propose the following problem as the linear dual problem (LDP) of our primal problem (LFP).

(LDP): Maximize $G(u) = \frac{b_o + c'u}{d_o}$

subject to $u \in L$

where

$$L = \{u : Au - b + \frac{1}{d_o}(b_o + c'u)d = 0,\ u \in -C^*,\ d_o \neq 0\}.$$

Following the traditional approach in establishing a "complete" duality theorem between (LFP) and (LDP), we begin by stating and proving a Weak Duality lemma.

<u>Lemma 4.2.1.</u> (Weak Duality). If $\bar{x}$ and $\bar{u}$ are feasible solutions of (LFP) and (LDP), respectively, then

$$f(\bar{x}) \geq G(\bar{u})$$

<u>Proof:</u> Since $\bar{u} \in L = \{u : Au - b + \frac{1}{d_o}(b_o + c'u)d = 0,\ u \in -C^*,$ $d_o \neq 0\}$ we have

$$A\bar{u} - b + \frac{(b_o + c'\bar{u})d}{d_o} = 0 ,$$

hence,

$$\bar{u}'A'\bar{x} - b'\bar{x} = - \frac{(b_o + c'\bar{u})d'\bar{x}}{d_o} .$$

Now $\bar{u}'A'\bar{x} \geq c'\bar{u}$,

since $\bar{u} \in -C^*$ and $c - A'x \in -C$,

which implies

$$c'\bar{u} - b'\bar{x} \leq - \frac{(b_o + c'\bar{u})d'\bar{x}}{d_o} . \tag{4.2.1}$$

Adding $-b_o$ to both sides of (4.2.1), we have that

$$c'\bar{u} - b'\bar{x} - b_o \leq \frac{(b_o+c'\bar{u})d'\bar{x}}{d_o} - b_o ,$$

$$-(b'\bar{x}+b_o) \leq - \frac{(b_o+c'\bar{u})d'\bar{x}}{d_o} - \frac{(b_o+c'\bar{u})d_o}{d_o} ,$$

$$-(b'\bar{x}+b_o) \leq - \frac{(b_o+c'\bar{u})(d'\bar{x}+d_o)}{d_o} . \qquad (4.2.2)$$

Dividing (4.2.2) by $-(d'\bar{x}+d_o)$,

$$\frac{b'\bar{x} + b_o}{d'x + d_o} \geq \frac{(b_o+c'\bar{u})}{d_o}$$

therefore,

$$f(\bar{x}) \geq G(\bar{u}) .$$

Lemma 4.2.2. If $\bar{x}$ and $\bar{u}$ are feasible solutions of (LFP) and (LDP), respectively, and $f(\bar{x}) = G(\bar{u})$, then $\bar{x}$ and $\bar{u}$ are optimal solutions.

Proof: Let $\bar{x}$ and $\bar{u}$ be feasible solutions, then appealing to Lemma 4.2.1, we have

$$f(\bar{x}) \geq G(\bar{u})$$

moreover,

$$f(x) \geq G(\bar{u})$$

for any x feasible solution of (LFP), and

$$f(x) \geq G(\bar{u}) = f(\bar{x}) ;$$

i.e. $\bar{x}$ solves (LFP) .

Also for any feasible vector u of (LDP) we have

$$f(\bar{x}) \geq G(u)$$

hence,

$$G(\bar{u}) = f(\bar{x}) \geq G(u) \;;$$

i.e. $\bar{u}$ solves (LDP) .

<u>Lemma 4.2.3.</u> For any feasible solution $(\bar{x},\bar{v})$ of (QFP), we have that

(i) $\bar{u} = \bar{v}\,(d'\bar{x}+d_o)$ is a feasible solution of (LDP),

and

(ii) $F(\bar{x},\bar{v}) = G(\bar{u})$.

<u>Proof:</u> To show first that $F(\bar{x},\bar{v}) = G(\bar{u})$. Let $(\bar{x},\bar{v}) \varepsilon D$, then

$$\nabla_x F(\bar{x},\bar{v}) = 0 \quad \text{and} \quad \bar{v} \varepsilon -C^*$$

or

$$\nabla_x \left(\frac{b'\bar{x} + b_o}{d'\bar{x} + d_o}\right) + \bar{v}'(c-A'\bar{x}) = 0 \;.$$

Hence, from (i)

$$\nabla_x \left(\frac{b'\bar{x} + b_o}{d'\bar{x} + d_o}\right) + \bar{u}' \left(\frac{c- A'\bar{x}}{d'\bar{x} + d_o}\right) = 0$$

with $\bar{u} \varepsilon -C^*$.

Therefore,

$$(d'\bar{x}+d_o)(b-A\bar{u}) = [b'\bar{x} + b_o + \bar{u}'(c-A'\bar{x})]d \;. \qquad (4.2.3)$$

Multiplying (4.2.3) by $\bar{x}$ we have

$$(d'\bar{x}+d_o)(b'\bar{x}-\bar{x}'A\bar{u}) = [b'\bar{x} + b_o + \bar{u}'(c-A'\bar{x})]d'\bar{x} \qquad (4.2.4)$$

and adding $d_o(b_o+c'\bar{u})$ to both sides of (4.2.4) gives

$$d_o(b'\bar{x} - \bar{x}'A\bar{u} + b_o + c'\bar{u}) = (d'\bar{x}+d_o)(b_o+c'\bar{u}) \ . \tag{4.2.5}$$

Now dividing (4.2.5) by d_o and $(d'\bar{x}+d_o)$ we have,

$$\frac{b'\bar{x} + b_o + \bar{u}'(c-A'\bar{x})}{d'\bar{x} + d_o} = \frac{b_o + c'\bar{u}}{d_o} \ ,$$

$$\frac{b'\bar{x} + b_o}{d'\bar{x} + d_o} + \bar{v}'(c-A'\bar{x}) = \frac{b_o + c'\bar{u}}{d_o} \ .$$

Hence,

$$F(\bar{x},\bar{v}) = G(\bar{u}) \ . \tag{4.2.6}$$

We show next that $\bar{u} \in L$. From (4.2.3) and with $(\bar{x},\bar{v}) \in D$ and $\bar{u} \in -C^*$

$$A\bar{u} - b + F(\bar{x},\bar{v})d = 0$$

and from (4.2.6) we have,

$$A\bar{u} - b + G(\bar{u})d = 0$$

or

$$A\bar{u} - b + \frac{(b_o+c'\bar{u})}{d_o}d = 0$$

Hence,

$$\bar{u} \in L \ .$$

Lemma 4.2.4. If for an arbitrary $\bar{x} \in P \subset E^n$ and an arbitrary $\bar{u} \in -C^* \subset E^m$ we have $F(\bar{x},\bar{v}) = G(\bar{u})$ where $\bar{v} = \bar{u}/(d'\bar{x}+d_o)$, then $(\bar{x},\bar{v}) \in D$ and $\bar{u} \in L$.

Proof: Since $F(\bar{x},\bar{v}) = G(\bar{u})$,

$$\frac{b'\bar{x} + b_o}{d'\bar{x} + d_o} + \bar{v}'(c-A'\bar{x}) = \frac{b_o + c'\bar{u}}{d_o}$$

and with $\bar{v} = \bar{u}/(d'\bar{x}+d_o)$ we have

$$\frac{b'\bar{x} + b_o + \bar{u}'(c-A'\bar{x})}{d'\bar{x} + d_o} = \frac{b_o + c'\bar{u}}{d_o} .$$

Therefore,

$$d_o[b'\bar{x} + b_o + \bar{u}'(c-A'\bar{x})] = (d'\bar{x}+d_o)(b_o+c'\bar{u}) .$$

This last equality is (4.2.5), now by reversing the steps from (4.2.3)-(4.2.5) with $\bar{x}$ arbitrary we have

$$(d'\bar{x}+d_o)(b-A\bar{u}) = [b'\bar{x} + b_o + \bar{u}'(c-A'\bar{x})]d \qquad (4.2.7)$$

or (4.2.3) which we have shown is equal to $\nabla_x F(\bar{x},\bar{v}) = 0$ where $\bar{v} = \bar{u}/(d'\bar{x}+d_o)$. Also, since $\bar{u} \in -C^*$ and $1/(d'\bar{x}+d_o) > 0$ we have $\bar{u}/(d'\bar{x}+d_o) \in -C^*$ which implies $\bar{v} \in -C^*$, therefore, $(\bar{x},\bar{v}) \in D$.

Rearranging (4.2.7) gives us

$$A\bar{u} - b + \left(\frac{b'\bar{x} + b_o + \bar{u}'(c-A'\bar{x})}{d'\bar{x} + d_o}\right) d = 0$$

or

$$A\bar{u} - b + \left(\frac{b'\bar{x} + b_o}{d'\bar{x} + d_o} + \bar{v}'(c-A'\bar{x})\right) d = 0 \quad ,$$

hence,

$$A\bar{u} - b + F(\bar{x},\bar{v})d = 0 \ .$$

Substituting $F(\bar{x},\bar{v})$ for $G(\bar{u})$,

$$A\bar{u} - b + G(\bar{u})d = 0 \ ,$$

$$A\bar{u} - b + \frac{(b_o+c'\bar{u})d}{d_o} = 0 \ ,$$

with $\bar{u} \in -C^*$, therefore, $\bar{u} \in L$.

We now establish the relationship between (LFP) and (LDP) by first proving the following Direct Duality theorem.

Theorem 4.2.1. (Direct Duality Theorem). If x^o solves (LFP), then there exists $u^o \in L$ which solves (LDP), and $f(x^o) = G(u^o)$.

Proof: If x^o solves (LFP), then from Theorem 3.2.2, there exists $v^o \in -C^*$ such that,

$$\nabla_x' f(x^o) + v^{o\prime}\nabla_x(c-A'x^o) = 0$$

$$v^{o\prime}(c-A'x^o) = 0$$

$$c - A'x^o \in C \ .$$

This gives us that $(x^o,v^o) \in D$, and $f(x^o) = F(x^o,v^o)$. Also, in view of Lemma 4.2.3, $F(x^o,v^o) = G(u^o)$, where $u^o = v^o(d'x^o+d_o) \in L$, therefore, from Lemma 4.2.2, u^o solves (LDP).

The converse duality relationship between (LDP) and (LFP) can be established through (QFP).

Lemma 4.2.5. Let $(x^o, v^o) \varepsilon D$. If $u^o = v^o(c'x^o + d_o)$ solves (LDP), then (x^o, v^o) solves (QFP).

Proof: Let u^o solve (LDP), and assume (x^o, v^o) does not solve (QFP); i.e. $F(x^o, v^o) < F(x^*, v^*)$ where (x^*, v^*) is the optimal solution of (QFP). Then from Lemma 4.2.3

$$F(x^*, v^*) = G(u^*)$$

where $u^* \varepsilon L$ and,

$$F(x^o, v^o) = G(u^o)$$

where $u^o \varepsilon L$.

In particular,

$$G(u^o) < G(u^*)$$

with u^o, $u^* \varepsilon L$, therefore, u^o does not solve (LDP). Hence, the result follows.

Lemma 4.2.6. If (x^o, v^o) solves (QFP), and the Hessian matrix of $F(x,v)$ is nonsingular at (x^o, v^o), then x^o solves (LFP).

Proof: If (x^o, v^o) solves (QFP), then from Theorem 3.2.2, there exists $w^o \varepsilon E^n$ which satisfies (a), (b), (c), (d), and (e) where

$$H(x,v,w) = f(x) + v'(c-Ax) + w'[\nabla_x f(x) + \left(\nabla_x v'(c-A'x)\right)'] ,$$

(a) $\nabla_x' f(x^o) + v^{o\prime}\nabla_x(c-A'x^o) + w^{o\prime}\nabla_x[\nabla_x f(x^o) + (v^{o\prime}\nabla_x(c-A'x^o))'] = 0$

(b) $(c-A'x^o) + [w^{o\prime}\nabla_v(v^{o\prime}\nabla_x(c-A'x^o))']' \in C$

(c) $\nabla_x f(x^o) + (v^{o\prime}\nabla_x(c-A'x^o))' = 0$

(d) $v^{o\prime}(c-A'x^o) + v^{o\prime}[w^{o\prime}\nabla_v(v^{o\prime}\nabla_x(c-A'x^o))']' = 0$

and

(e) $v^o \in -C^*$.

In view of (c), (a) can be reduced to the following

$$w^{o\prime}\nabla_x[\nabla_x f(x^o) + (v^{o\prime}\nabla_x(c-A'x^o))'] = 0$$

$$w^{o\prime}\nabla_x^2 F(x^o,v^o) = 0 \ .$$

However, by assumption $\nabla_x^2 F(x^o,v^o)$ is nonsingular implying that $w^o = 0$; hence, (a) can be reduced to

$$\nabla_x f(x^o) + v^{o\prime}\nabla_x(c-A'x^o) = 0 \tag{4.2.8}$$

Also, since $w^o = 0$, then (d) and (b) can be written as

$$v^{o\prime}(c-Ax^o) = 0 \tag{4.2.9}$$

and

$$c - Ax^o \in C \, . \tag{4.2.10}$$

In view of (4.2.8)-(4.2.10) and (e) we apply Theorem 3.2.3, and it follows that x^o solves (LFP).

In Lemma 4.2.5 we assume that (x^o,u^o) was a feasible solution of (QFP). A stronger converse relation can be established between (LFP) and (LDP) by paralleling the arguments of Charnes and Cooper (1962).

Employing the transformation

$$y = \rho x,$$
$$\rho \geq 0$$

which is a homeomorphisin, we shall show that an optimal solution of (LFP) can be obtained by solving the following equivalent linear program.

(ELP): Minimize $\psi(y,\rho) = b'y + b_o\rho$

subject to $(y,\rho) \in P_\rho \subset E^{n+1}$

where

$$P_\rho = \{(y,\rho) : A'y - c\rho \in -C,\ d'y + d_o\rho = 1,$$
$$\rho \in \text{Int } Q_1^+,\ y \in E^n\} .$$

<u>Lemma 4.2.7.</u> Every (y,ρ) satisfying the constraints of P_ρ has $\rho > 0$.

<u>Proof:</u> Suppose $(\hat{y},0)$ satisfy the constraints of P_ρ . Let $\hat{x}$ be any element of X . Then $x_u = \hat{x} + u\hat{y}$ is in X for $\hat{u} > 0$ since $-A\hat{y} \in C$, $\hat{y} \in E^n$. But then X is unbounded contrary to the regularity hypothesis (vi) imposed on X in (LFP).

<u>Lemma 4.2.8.</u> If (y^o,ρ^o) is an optimal solution of (ELP), then $x^o = y^o/\rho^o$ is an optimal solution of (LFP).

<u>Proof:</u> Suppose the theorem is false; i.e. assume that there exists an optimal $x^* \in X$ such that

$$\frac{b'x^* + b_o}{d'x^* + d_o} < \frac{b'(y^o/\rho^o) + b_o}{d'(y^o/\rho^o) + d_o} .$$

Since $d'x^* + d_o > 0$ we have that

$$d'x^* + d_o = \theta \cdot 1$$

for some $\theta > 0$.

Consider $\hat{y} = \theta^{-1} x^*$, and let $\hat{\rho} = \theta^{-1}$.

Then

$$\theta^{-1}(d'x^* + d_o) = d'\hat{y} + d_o\hat{\rho} = 1$$

and $(\hat{y},\hat{\rho})$ also satisfies,

$$A'\hat{y} - c\hat{\rho} \in -C,$$

$$y \in E^n,$$

$$\hat{\rho} \in \text{Int } Q_1^+ .$$

Now,

$$\frac{b'x^* + b_o}{d'x^* + d_o} = \frac{\theta^{-1}(b'x^* + b_o)}{\theta^{-1}(d'x^* + d_o)} = \frac{b'\hat{y} + b_o\hat{\rho}}{d'\hat{y} + d_o\hat{\rho}} = \frac{b'\hat{y} + b_o\hat{\rho}}{1} .$$

Also,

$$\frac{b'(y^o/\rho^o) + b_o}{d'(y^o/\rho^o) + d_o} = \frac{b'y^o + b_o\rho^o}{d'y^o + d_o\rho^o} = \frac{b'y^o + b_o\rho^o}{1} .$$

This implies,

$$b'\hat{y} + b_o\hat{\rho} < b'y^o + b_o\rho^o$$

which contradicts our assumption that (y^o,ρ^o) is an optimal solution of (ELP).

Lemma 4.2.9. (LDP) is a dual problem of (ELP).

Proof: (LDP) is equivalent to the following equivalent linear dual program (ELDP) obtained by substituting $\frac{b_o + c'u}{d_o} = z$ in (LDP).

(ELDP): Maximize z

$$\begin{aligned} \text{subject to} \quad & Au + dz = b \\ & -c'u + d_o z \le b_o \\ & u \in -C^* \\ & z \in E^1 . \end{aligned}$$

By Ben-Israel (1969b) we have that (ELDP) is a linear dual problem to (ELP), since

(ELDP): Maximize $(0,1)'\begin{pmatrix}1\\ z\end{pmatrix}$

$$\text{subject to} \quad \begin{pmatrix} A & d \\ -c' & d_o \end{pmatrix}\begin{pmatrix} u \\ z \end{pmatrix} \begin{matrix} = \\ \le \end{matrix} \begin{pmatrix} b \\ b_o \end{pmatrix}$$

$$u \in -C^*$$

$$z \in E^1$$

and

(ELP): Minimize $(b,b_o)'\begin{pmatrix}y\\ \rho\end{pmatrix}$

$$\begin{pmatrix} A' & -c \\ d' & d_o \end{pmatrix}\begin{pmatrix} y \\ \rho \end{pmatrix} - \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{matrix} \in -C \\ \in \{0\} \end{matrix}$$

$$y \in E^n$$

$$\rho \in Q_1^+ .$$

Using Lemma 4.2.8 we can prove the Converse Duality theorem between (LFP) and (LDP).

Theorem 4.2.2. (Converse Duality theorem). If $u^o \in L$ is an optimal solution to the (LDP), then there exists an $x^o \in X \subset E^n$, which is an optimal solution to problem (LFP) and $G(u^o) = f(x^o)$.

Proof: As a consequence of Lemma 4.2.6 and well known results in linear duality, it follows that, if u^o is an optimal solution of (LDP), then there exists a (y^o,ρ^o) which is an optimal solution of (ELP). Furthermore, by Lemma 4.2.8 $x^o = y^o/\rho^o$, exists and is an optimal solution of (LFP). Thus, once u^o is known, (y^o,ρ^o) exists and if it could be computed say via the Simplex method, x^o could be computed.

We can provide an alternative method of computing the optimal solution of (LFP) and (QFP) by solving (LDP).

Theorem 4.2.3. If $u^o \in L$ is an optimal solution to (LDP), then, there exists an $x^o \in X \subset E^n$ such that $(x^o,v^o) \in D$ is an optimal solution of (QFP) where $v^o = u^o/(d'x^o+d_o)$.

Proof: Let $u^o \in L$ be an optimal solution to (LDP). Therefore, by the Converse Duality theorem between (LFP) and (LDP), there exists an $x^o \in X$ which optimizes $f(x)$, and

$$G(u^o) = f(x^o) \ . \tag{4.2.11}$$

Since $u^o \in L$,

$$Au^o - b + \frac{1}{d_o}(b_o+c'u^o)d = 0 \ .$$

From (4.2.11) we have that

$$\frac{b_o + c'u^o}{d_o} = \frac{b'x^o + b_o}{d'x^o + d_o}$$

this implies

$$Au^o - b + \left(\frac{b'x^o + b_o}{d'x^o + d_o}\right)d = 0$$

multiplying by $\dfrac{1}{d'x^o + d_o} \neq 0$ we have that

$$\frac{Au^o}{d'x^o + d_o} - \frac{b}{d'x^o + d_o} + \frac{(b'x^o+b_o)d}{(d'x^o+d_o)^2} = 0 \ ,$$

$$-\ \frac{[(d'x^o+d_o)b - (b'x^o+b_o)d]}{(d'x^o+d_o)^2} + \frac{Au^o}{d'x^o + d_o} = 0 \ ,$$

$$-\ \nabla_x f(x^o) + Av^o = 0 \ ,$$

$$-\ \nabla_x f(x^o) - v^{o\prime}\nabla_x(c-A'x^o) = 0 \ ,$$

$$\nabla_x f(x^o) + v^{o\prime}\nabla_x(c-A'x^o) = 0 \ ,$$

$$\nabla_x F(x^o,v^o) = 0 \ .$$

Therefore,

$$(x^o,v^o) \varepsilon D \ .$$

By Lemma 4.2.5, $(x^o,v^o) \varepsilon D$ implies $F(x^o,v^o) = G(u^o)$. If $(x^o,v^o) \varepsilon D$ does not optimize $F(x,v)$, then let $(x^*,v^*) \varepsilon D$ be a global maximum of $F(x,v)$ on D .

This implies that $u^* \in L$, and

$$G(u^*) = F(x^*,v^*) > F(x^o,v^o) = G(u^o)$$

which is a contradiction.

A summary of the relationships established in this chapter is given in Figure 4.1.

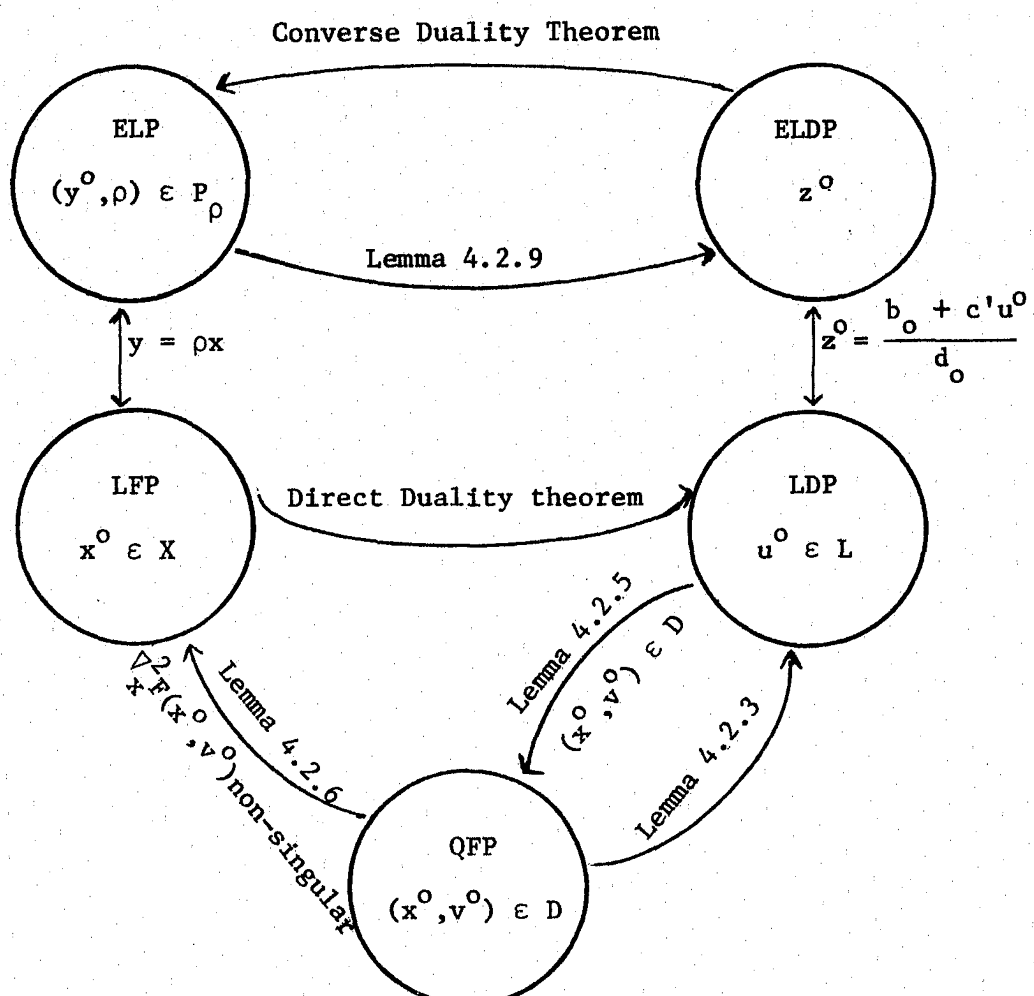

Figure 4.1. Dual Relationships of Fractional Problems

5. BIBLIOGRAPHY

Abrams, R. A. 1973. Nonlinear programming in complex space: sufficient conditions and duality. Journal of Mathematical Analysis and Applications 38:619-32.

Abrams, R. A. and A. Ben-Israel. 1969. A duality theorem for complex quadratic programming. Journal of Optimization Theory and Applications 4(4):244-52.

Alders, C. P. 1976. Generalized convex programming over cone domains. Ph.D. dissertation. Library, Iowa State University, Ames, Iowa.

Alders, C. P. and V. A. Sposito. 1978. Necessary optimality conditions for programming problems over cone domains. Math Operationsforsh Statist., Ser. Optimization 9(4):515-21.

Bector, C. R. 1971a. Duality in linear fractional programming. Research Report No. 4. Faculty of Commerce, University of Manitoba, Winnipeg 19, Canada.

Bector, C. R. 1971b. Duality in non-linear fractional programming. Research Report No. 5. Faculty of Commerce, University of Manitoba, Winnipeg 19, Canada.

Bector, C. R. 1973a. Duality in linear fractional programming. Utilitas Mathematica 4:155-68.

Bector, C. R. 1973b. Duality in nonlinear fractional programming. Zeitschrift für Operations Research 17:183-93.

Bector, C. R. 1974. On convexity pseudo-convexity and quasi-convexity of composite functions. Cahiers du Centre d' Études de Recherche Operationnelle 15(1):411-28.

Bector, C. R. and S. K. Bhatt. 1978. Pseudo-monotonic interval programming. Naval Research Logistics Quarterly 25(2):309-14.

Bector, C. R., K. Bector, and J. E. Klassen. 1977. Duality for a nonlinear programming problem. Utilitas Mathematica 11:87-99.

Ben-Israel, A. 1969a. Theorems of the alternative for complex linear inequalities. Israel Journal of Mathematics 7:129-36.

Ben-Israel, A. 1969b. Linear equations and inequalities on finite dimensional, real or complex, vector spaces: a unified theory. Journal of Mathematical Analysis and Applications 27:367-89.

Berman, A. and A. Ben-Israel. 1971. More on linear inequalities with applications to matrix theory. Journal of Mathematical Analysis and Applications 33(3):482-96.

Bhatt, S. K. and S. K. Misra. 1975. Sufficient optimality criteria in non-linear programming in the presence of convex equality and inequality constraints. Zeitschrift für Operations Research 19:101-05.

Bourbaki, N. 1953. Espaces vectoriels topologiques. Hermann and Cie, Paris.

Chadha, S. S. and S. Shivpuri. 1980. Enumerative technique for an extreme point fractional program. European Journal of Operational Research 4(1):54-59.

Chandra, S. and T. R. Gulati. 1976. Duality theorem for a non-differentiable fractional programming problem. Management Science 23(1):32-37.

Charnes, A. and W. W. Cooper. 1962. Programming with linear fractional functionals. Naval Research Logistic Quarterly 9(3-4):181-86.

Charnes, A. and W. W. Cooper. 1973. An explicit general solution in linear fractional programming. Naval Research Logistic Quarterly 20(3):449-67.

Charnes, A., D. Granot, and F. Granot. 1976. A note on explicit solution in linear fractional programming. Naval Research Logistic Quarterly 23(1):161-67.

Craven, B. D. and B. Mond. 1973. The dual of a fractional linear program. Journal of Mathematical Analysis and Application 42:507-12.

Craven, B. D., and B. Mond. 1975. On fractional programming and equivalence. Naval Research Logistic Quarterly 22:405-10.

Dantzig, G. B. 1963. Linear programming and extensions. Princeton University Press, Princeton, NJ.

Dinkelbach, W. 1967. On nonlinear fractional programming. Management Science 13(7):492-98.

Guignard, M. 1969. Generalized Kuhn-Tucker conditions for mathematical programming problems in a Banach space. SIAM Journal on Control Theory 7:232-41.

Hadley, G. 1964. Nonlinear and Dynamic Programming. Addison-Wesley, Cambridge, Massachusetts.

Hanson, M. A. 1961. Duality theorem in non-linear programming with non-linear constraints. Australian Journal of Statistics 3:64-71.

Kuhn, H. W., and A. W. Tucker. 1951. Nonlinear programming. Pages 481-492 in Second Berkeley Symposium Proceedings on Mathematical Statistics and Probability. University of California Press, Berkeley, California.

Luenberger, D. G. 1969. Optimization By Vector Space Methods. Wiley, New York.

Mangasarian, O. L. 1969. Nonlinear Programming. McGraw-Hill, New York.

Nikaidô, H. 1954. On von Neumann's minimax theorem. Pacific Journal of Mathematics 4:65-72.

Ritter, K. 1967. Duality for nonlinear programming in a Banach space. SIAM Journal of Applied Mathematics 15(2):294-302.

Schaible, S. 1976. Duality in fractional programming: a unified approach. Operation Research 24(3):452-61.

Sharma, I. C. and K. Swarup. 1972. On duality in linear fractional functionals programming. Zeitschrift für Operation Research 16:91-100.

Singh, C. 1977. Sufficient optimality criteria in nonlinear programming for generalized equality-inequality constraints. Journal of Optimization Theory and Application 22(4):631-35.

Skarpness, B. O., and V. A. Sposito. 1980. A modified fritz john optimality criterion. Journal of Optimization Theory and Application 31(1):113-15.

Slater, M. 1951. Lagrange multipliers revisited: a contribution to nonlinear programming. Rand Corporation Report RM-676, Santa Monica, California.

Sposito, V. A. 1974. Modified regularity conditions for nonlinear programming problems over mixed cone domains. Journal of Mathematical Programming 6:167-79.

Sposito, V. A. 1975. Linear and Nonlinear Programming. Iowa State University Press, Ames, Iowa.

Sposito, V. A. 1976. Quadratic duality over convex cone domains. Journal of Mathematical Programming 10:277-83.

Sposito, V. A., and H. T. David. 1971. Saddlepoint optimality criteria of nonlinear programming problems over cones without differentiability. SIAM Journal of Applied Mathematics 20(4):698-702.

Sposito, V. A., and H. T. David. 1972. A note on Farkas lemmas over cone domains. SIAM Journal of Mathematics 22(3):356-58.

Swarup, K. 1968. Duality in fractional programming. Untunchmensforschung 12(2):106-12.

Tuy, Hoàng. 1964. Sur les inégalités linéaires. Colloquium Mathematician 13:107-23.

Vajda, S. 1972. Probablistic Programming. Academic Press, New York.

Varaiya, P. O. 1967. Nonlinear programming in Banach space. SIAM Journal on Applied Mathematics 15:284-93.

6. ACKNOWLEDGMENTS

I would like to express my thanks and deepest appreciation to Professor Vincent A. Sposito. Because of Professor Sposito's continued encouragement, guidance, and time it was possible for me to complete this endeavour, and it was a privilege to work with a person with such enthusiasm, insight, and knowledge in the area of mathematical programming. I am also grateful to him for his contribution to my education and for the friendship which developed during our association, a friendship which I deeply value.

I would also like to express my appreciation to my co-major professor, Dr. Keith L. McRoberts, for his guidance and time in directing my studies in Industrial Engineering.

I wish to thank my other committee members, Dr. William J. Kennedy, Dr. Howard D. Meeks, Dr. Chien-Pai Han, and Dr. Clair G. Maple, for their willingness to serve on the committee.

Many thanks to my colleague and good friend, Luis A. Escobar, for the many hours of fruitful discussion.

Finally, I would like to thank my parents, Marlene and Norman Skarpness, for the continuous support they have given me during my many years of graduate study.